TESE DE DOUTORAMENTO

EFFICIENT REGISTRATION OF MULTI AND HYPERSPECTRAL REMOTE SENSING IMAGES ON GPU

Álvaro Ordóñez Iglesias

ESCOLA DE DOUTORAMENTO INTERNACIONAL DA UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOUTORAMENTO EN INVESTIGACIÓN EN TECNOLOXÍAS DA INFORMACIÓN

SANTIAGO DE COMPOSTELA
2021

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Declaración do autor da tese
EFFICIENT REGISTRATION OF MULTI AND HYPERSPECTRAL REMOTE SENSING IMAGES ON GPU

Don Álvaro Ordóñez Iglesias

Presento a miña tese, seguindo o procedemento axeitado ao Regulamento, e declaro que:

1. *A tese abarca os resultados da elaboración do meu traballo.*
2. *De ser o caso, na tese faise referencia ás colaboracións que tivo este traballo.*
3. *Confirmo que a tese non incorre en ningún tipo de plaxio doutros autores nin de traballos presentados por min para a obtención doutros títulos.*
4. *A tese é a versión definitiva presentada para a súa defensa e coincide a versión impresa coa presentada en formato electrónico.*

E comprométome a presentar o Compromiso Documental de Supervisión no caso de que o orixinal non estea na Escola.

En Santiago de Compostela, 19 de xullo de 2021

Asdo. Álvaro Ordóñez Iglesias

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Autorización do director/titor da tese
EFFICIENT REGISTRATION OF MULTI AND HYPERSPECTRAL REMOTE SENSING IMAGES ON GPU

Dona Dora Blanco Heras, Profesora Titular da Área de Arquitectura e Tecnoloxías de Computadores da Universidade de Santiago de Compostela

Don Francisco Santiago Argüello Pedreira, Profesor Titular da Área de Arquitectura e Tecnoloxías de Computadores da Universidade de Santiago de Compostela

INFORMAN:

*Que a presente tese correspóndese co traballo realizado por **Don Álvaro Ordóñez Iglesias**, baixo a nosa dirección/titorización, e autorizamos a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como directores/titores desta non incorre nas causas de abstención establecidas na Lei 40/2015.*

De acordo co indicado no Regulamento de Estudos de Doutoramento, declaramos tamén que a presente tese de doutoramento é idónea para ser defendida en base á modalidade de COMPENDIO DE PUBLICACIÓNS, nos que a participación do doutorando foi decisiva para a súa elaboración e as publicacións se axustan ao Plan de Investigación.

En Santiago de Compostela, 19 de xullo de 2021

Asdo. Dora Blanco Heras
Directora tese

Asdo. Francisco Santiago Argüello Pedreira
Director tese

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Perhaps when distant people on other planets pick up some wavelength of ours all they hear is a continuous scream.

Iris Murdoch

A loita nobre e leal das ideias é o que asegura o progreso.

Alfonso Daniel Rodríguez Castelao

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Agradecementos

Esta sección non podería comezar doutra maneira que expresando o maior dos meus agradecementos aos meus directores Dora e Francisco pola confianza depositada e por todo o seu apoio, titorización e guía ao longo destes anos.

Darlle as grazas ao Departamento de Electrónica e Computación, así como ao Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), ambos da Universidade de Santiago de Compostela, por brindarme as ferramentas e os espazos necesarios para a realización deste traballo. Agradecer tamén a labor das unidades técnicas e de apoio do CiTIUS pola súa dispoñibilidade en todos os asuntos burocráticos e tecnolóxicos que se me foron cruzando nesta etapa.

My gratitude also goes to Prof. Dr. Begüm Demir and her group, the Remote Sensing Image Analysis group of the Technische Universität Berlin, for their advise and kind support during my research stay in Berlin. I would also like to thank the people I met there who contributed to making my stay an unforgettable professional and personal experience.

Non quero esquecerme dos meus compañeiros e compañeiras do CiTIUS, que me axudaron a adozar o día a día cun café, cun xantar, cunha cea, ou simplemente parolando: Adrián, Ángel, Celia, Eric, Fernando Estévez, Fernando Gómez, Jorge, Lorenzo, Nico, Rubén, Yago... En especial, ao compañeiro de andainas, David Chapela, que case case fomos un turrando da man do outro; ao compañeiro de habitáculo e cervexas, Víctor; e á da tintura menos tintura que vin nunca, Andrea. Grazas tamén aos compañeiros do equipo hiperespectral: Pablo, Javi, Alberto, Álvaro Acción, Pedro e Sergio. Tamén a Richard por toda a axuda brindada.

Mención especial para as compañeiras da Asemblea de Investigadoras de Compostela por crear un espazo moi necesario na nosa universidade, e por loitar pola dignificación e os dereitos de todas as investigadoras.

O meu agradecemento vai tamén para as entidades que financiaron este traballo. Esta tese

doutoral foi financiada polo Ministerio de Universidades mediante un contrato a cargo do programa de Formación de Profesorado Universitario (FPU16/03537 e EST18/00602) e mediante diferentes proxectos financiados polo Ministerio de Ciencia e Innovación (TIN2013-41129-P, TIN2016-76373-P e PID2019-104834GB-I00) e pola Consellería de Cultura, Educación e Universidade (GRC2014/008, ED431G/08, ED431C 2018/19 e ED431G-2019/04) co cofinanciamento do Fondo Europeo de Desenvolvemento Rexional (FEDER).

Por último, pero non menos importante, darlle as grazas á miña familia, por todo os sacrificios que fixeron por min e porque sen eles hoxe non sería quen son. Tamén, á persoa que máis tempo lle roubei ("*Poño unhas probas a correr, 5 minutos e xa saio de traballar*" e poucas veces eran 5 minutos), á miña compañeira Raquel, e que moito lle teño rompido a cabeza durante estes anos.

19 de xullo de 2021

Resumo

O grande avance no desenvolvemento de sensores permite que hoxe poidamos obter imaxes multiespectrais e hiperespectrais de teledetección a menores custos que antano. A característica máis significativa destas imaxes é o rango de lonxitude de onda que abarcan [13]. Dende o visible, pasando polo infravermello próximo e ata o infravermello medio do espectro electromagnético, é dicir, dende 0.4 a 2.5 µm [34]. As imaxes multi e hiperespectrais están formadas por moitas canles espectrais continuas, tamén chamadas bandas. Unha imaxe con 10 ou menos bandas recibe o nome de imaxe multiespectral, mentres que unha imaxe con 10 ou máis bandas, pero habitualmente máis de 100, recibe o nome de imaxe hiperespectral. Isto significa que cada píxel é un vector onde cada compoñente se corresponde con unha determinada lonxitude de onda [49]. Grazas a alta resolución espectral destas imaxes podemos distinguir con maior precisión obxectos, especies, usos do chan, entre outros.

Centos de imaxes hiperespectrais son obtidas todos os días grazas ao crecente número de satélites e aeronaves con sensores espectrais [20, 25, 61]. Isto permítenos obter imaxes da mesma zona da Terra tomadas en distintos instantes de tempo (minutos, horas, meses ou incluso anos) e dende diferentes perspectivas. Esta serie de imaxes é empregada en tarefas onde é esencial comparar, estudar ou buscar as diferenzas entre imaxes. Algúns exemplos destas tarefas ou aplicacións son a detección automática de cambios [32, 59, 77, 79], o control medioambiental [40, 120, 138], a fusión de imaxes [7, 14, 63, 87] ou a creación de imaxes con superresolución [4, 68, 69]. Antes de levar a cabo calquera procesado asociado a ditas tarefas, as imaxes teñen que ser rexistradas.

O rexistro, aliñamento ou superposición de imaxes consiste en aliñar imaxes da mesma zona ou área pero que foron obtidas en distintos instantes de tempo, dende diferentes perspectivas, e posiblemente baixo diferentes condicións de luz. Se nos limitamos ao rexistro de dúas imaxes, o que se pretende é rexistrar ou aliñar unha delas, chamada imaxe de referencia, con respecto

da outra, chamada imaxe obxectivo. Isto esixe localizar e emparellar en ambas imaxes os mesmos obxectos, estruturas, rexións, etc. O rexistro non é doado pois habitualmente as imaxes presentan cambios debido ao paso do tempo dende que foi tomada cada unha delas, mais tamén presentan cambios de iluminación, de resolución espacial e espectral, etc., e pode que non cubran exactamente a mesma rexión.

A todo isto hai que sumarlle outras eivas asociadas ás imaxes de teledetección. Estas imaxes presentan certas características que fan do rexistro un tarefa difícil e complexa en comparación con outros tipos de imaxes, por exemplo, as imaxes médicas [71]. Os algoritmos de rexistro teñen que lidar coa ampla variedade de sensores e coas distintas condicións na que as imaxes poden ser obtidas; coas condicións atmosféricas e meteorolóxicas; coa estación do ano na que nos atopemos xa que poden facer que elementos naturais como árbores, lagoas e ríos sexan puntos de referencia nada ou pouco fiables para o rexistro; coas diferentes condicións de luz debido a efectos do paso do tempo, por exemplo, o cambio do ángulo do Sol durante o ano; e a grande cantidade de datos. Todo isto fai do rexistro de imaxes de teledetección terrestre unha difícil tarefa.

O rexistro de imaxes pode facerse á man. Habitualmente faise seleccionando puntos característicos (puntos de control) comúns en todas as imaxes a rexistrar. O rexistro manual non é viable cando falamos de imaxes de teledetección e máis especificamente cando falamos de imaxes tomadas por satélites, os cales, capturan centos de imaxes ao día. O proceso esixiría un alto custo temporal e de persoal para levalo a cabo. Ademais, o rexistro manual non é doado pois moitas veces é difícil atopar puntos de referencia debido á propia resolución das imaxes ou porque se tratan de imaxes de urbes ou zonas agrícolas con moita monotonía. Nace deste xeito o rexistro automático como unha área de investigación dentro do procesamento de imaxes.

Os métodos de rexistro automático de imaxes poden clasificarse en dúas categorías [149]:

- **Métodos baseados en área**. Estes métodos fundaméntanse no cómputo das diferenzas entre píxeles [71]. Os métodos baseados en área explotan directamente os datos das imaxes sen ningunha análise estrutural, é dicir, non buscan correspondencias entre as imaxes procurando puntos ou rexións comúns. Estes métodos poden á súa vez dividirse en tres categorías: métodos baseados en información mutua [15, 65], métodos baseados na correlación [9, 62] e métodos baseados na transformada de Fourier [16, 122]. Os métodos baseados en información mutua precisan que a intensidade entre as imaxes sexa estable, o cal non é unha característica moi común nas imaxes de teledetección

terrestre. Por esta razón, estes métodos son habitualmente empregados en imaxe médica onde as condicións para a obtención das imaxes está moi controlada ao seren adquiridas en laboratorios ou hospitais. Os métodos da segunda categoría, os métodos baseados na correlación, baséanse en optimizar unha medida de semellanza desprazando rexións da imaxe de referencia á imaxe obxectivo. A medida de semellanza compútase para pares de ventás de ambas imaxes o que fai que este tipo de métodos sexan computacionalmente caros.

Polo contrario, os métodos baseados na transformada de Fourier son unha alternativa moi eficiente en termos computacionais para calcular a correlación entre imaxes. Estes métodos empregan o teorema do desprazamento para a transformada de Fourier o cal indica que un desprazamento circular no dominio espacial equivale a unha diferenza de fase no dominio da frecuencia. Este tipo de métodos son unha boa escolla cando se trata de rexistrar imaxes que foron obtidas baixo diferentes condicións [71, 149]. Estas características fan dos métodos baseados na transformada de Fourier nunha opción interesante para o rexistro de imaxes hiperespectrais de teledetección terrestre. En xeral, os métodos baseados en área son computacionalmente moi eficientes pero a súa precisión pode verse afectada por oclusións, cambios temporais, ruído, ou outros efectos que poden producir cambios na intensidade dos píxeles.

- **Métodos baseados en características**. Este tipo de métodos procuran por características distintivas ou de interese nas imaxes a rexistrar. Estas características poden ser puntos, contornos, rexións ou liñas que cumpren cunha serie de requisitos: ser invariantes a transformacións xeométricas, fáciles de localizar en termos de precisión e ser insensitivas a degradacións. Estes métodos baséanse en detectar as mesmas características nas imaxes a rexistrar. Desta maneira, coñecendo un certo número de características comúns, pódese calcular unha transformación xeométrica que rexistre unha imaxe con respecto da outra. Esta procura de características comúns a alto nivel fai que este tipo de métodos se vexan menos afectados por cambios de iluminación, de intensidade pola presenza de ruído ou por tratarse dun rexistro de imaxes capturadas por diferentes sensores [53]. Ademais, as imaxes de teledetección adoitan conter moitas liñas, esquinas, rexións, etc., susceptibles de ser detectadas por este tipo de métodos. Estas razóns converten aos métodos de rexistro baseados en características nos máis empregados en canto a imaxes de teledetección.

  Pola contra, estes métodos requiren da construción dunha escala-espazo para detectar

características cos requisitos descritos, así como para describir cada unha delas mediante un vector único pero que non se vexa afectado por transformacións xeométricas, cambios de luz ou ruído. Isto tradúcese nun alto custo computacional.

Na categoría de métodos baseados en área destaca o algoritmo *Fourier-Mellin invariant symmetric phase-only matched filtering (FMI-SPOMF)* [16] mentres que nos métodos baseados en características destaca *Scale-Invariant Feature Transform (SIFT)* [80]. Estes algoritmos foron desenvolvidos para procesar imaxes en escala de grises ou imaxes *RGB* (do inglés, *Red Blue Green*) e non para lidar coa información espectral dispoñible nas imaxes multi e hiperespectrais. A información espectral destas imaxes bríndanos unha información extra que pode ser empregada para mellorar a precisión do rexistro. Hai características que están só presentes nalgunha das bandas das imaxes e que poden ser empregadas a hora de computar a transformación xeométrica que buscamos para aliñar as imaxes. Ademais, a información espectral, xunto coa información espacial, pode ser empregada para esa distinción inequívoca de se un punto ou rexión dunha imaxe é ou non é o equivalente noutra imaxe. Ao mesmo tempo, as imaxes hiperespectrais tamén contan con moita información redundante polo que é preciso efectuar unha selección da información relevante.

O uso da información espectral parece máis que xustificado mais ten unha eiva, o alto custo computacional do seu procesamento. O rexistro, ao igual que calquera outra tarefa habitual de procesado de imaxes multi e hiperespectrais (detección de cambios, clasificación, detección de obxectos, etc.), require dun alto custo temporal. Algunhas propostas da literatura [58, 74, 85, 132, 143] resolven esta eiva procesando as imaxes multi e hiperspectrais como imaxes dunha soa banda perdendo desta maneira todo o potencial deste tipo de imaxes e renunciando así a empregar estes datos para realizar rexistros máis precisos ou capaces de aliñar imaxes con maior diferenza de escala, rotación ou translación.

Noutros traballos [24, 88] propuxéronse métodos de rexistro tendo en conta a información espectral, mais ningún deles resolve o problema de rexistro eficiente de grandes conxuntos de imaxes hiperespectrais ou a súa realización en tempo real. O tempo de cómputo é crucial en moitos casos de uso, por exemplo, no control de desastres medioambientais. Ademais, o número de imaxes hiperespectrais nas bases de datos das axencias espaciais non deixa de medrar cada día que pasa, evidenciando a necesidade de métodos eficientes para procesar esa grande cantidade de datos. É tamén unha realidade que o uso deste tipo de imaxes se está estendendo a ámbitos máis humildes como o caso dos vehículos aéreos non tripulados (en inglés, *unmanned aerial vehicles (UAVs)*) grazas ao abaratamento dos prezos dos espectrómetros.

Todo isto fai indispensable o desenvolvemento de algoritmos de rexistro automáticos que sexan eficientes tanto aproveitando e explotando a información espectral para mellorar a calidade do rexistro como dende o punto de vista computacional.

É por todo isto que se definen como hipóteses desta tese as seguintes:

H1. *O rexistro automático de imaxes multi e hiperespectrais con grandes diferenzas a nivel escala e orientación, ademais de cambios, por exemplo, froito do paso do tempo, é posible grazas á explotación eficiente da información espectral.*

H2. *O rexistro automático de imaxes multi e hiperespectrais pode ser realizado en tempo real grazas ao desenvolvemento de algoritmos paralelos en hardware de baixo custo.*

O proceso de demostración desas hipóteses dan lugar o obxectivo principal desta tese. Este é o desenvolvemento de algoritmos eficientes de rexistro automático de imaxes multi e hiperespectrais de teledetección de cobertura terrestre. Para iso, os algoritmos desenvolvidos empregan a grande cantidade de información que estas imaxes conteñen, tanto de tipo espectral como espacial, conseguindo aliñar imaxes con escalas, rotacións e translacións superiores con respecto a traballos publicados na literatura.

Nesta tese doutoral proponse un método baseado na transformada de Fourier así como un esquema de rexistro adaptable a calquera detector e descritor de características. Este esquema foi a base de desenvolvemento de tres algoritmos de rexistro de imaxes hiperespectrais baseados en detección de características. Tamén se desenvolveron métodos específicos para a selección de bandas e o cómputo de transformación xeométrica a partir da extracción de características. Ambos forman parte do mencionado esquema de rexistro.

Ademais, realizáronse implementacións eficientes en unidades de procesamento gráfico (en inglés, *Graphics Processing Unit (GPU)*) que nos permiten rexistrar imaxes en tempo real así como grandes conxuntos de datos empregando *hardware* de baixo custo, é dicir, sen precisar dunha infraestrutura de alto rendemento como son os supercomputadores. Preséntase tamén unha implementación multi-GPU para a realización dun rexistro multinivel en paralelo de diferentes conxuntos de imaxes multiespectrais.

Para realizar a experimentación, construíuse un conxunto de datos de proba que está dispoñible para toda a comunidade científica [115]. Este conxunto está formado por pares de imaxes tomadas polos sensores ROSIS e AVIRIS en distintas datas, dende distintas perspectivas e que presentan cambios froito do paso do tempo e de iluminación así como distorsións. Os algoritmos e métodos desenvolvidos nesta tese foron comparados con métodos do estado da

arte en termos de precisión e rendemento obtendo bos resultados. Conseguiuse rexistrar imaxes con escalados de ata $24\times$ nun tempo de execución de entre 1 e 4 segundos de media.

Esta tese doutoral dou lugar a diversas publicacións en revistas e congresos internacionais. As máis relevantes son recollidas nos seguintes capítulos desta tese:

- No Capítulo 3 proponse Hyperspectral Fourier–Mellin (HYFM), un método para o rexistro de dúas imaxes hiperespectrais baseado na transformada de Fourier, os métodos de área máis eficientes. O método emprega a información contida nas diferentes bandas ao longo das súas seis etapas. A transformada rápida de Fourier (en inglés, *fast Fourier transform (FFT)*) é utilizada para o cómputo eficiente da correlación das dúas imaxes. Faise uso da transformada fraccional de Fourier (en inglés, *fractional Fourier transform (FRFT)*) e da transformada fraccional multicapa de Fourier (en inglés, *multi-layer fractional Fourier transform (MLFFT)*) para obter rexistros máis precisos pois redúcense os erros de interpolación grazas ao cómputo dunha grella multinivel. O método computa a correlación en oito pares de compoñentes principais calculadas na primeira etapa empregando a técnica de análise de compoñentes principais (en inglés, *Principal Component Analysis (PCA)*). Esta información é logo combinada na quinta etapa para así ter en conta toda a información espectral a hora de computar a transformación xeométrica que rexistre as imaxes. O método é comparado con outros métodos na literatura así como outros métodos de extracción de características distintos da análise de compoñentes principais.

- No Capítulo 4 preséntase unha implementación eficiente en GPU do método HYFM para o rexistro de imaxes hiperspectrais. O método é executado totalmente en GPU explotando así súa arquitectura de cómputo e evitando as transferencias de datos innecesarios entre a CPU e a GPU que lastrarían o rendemento. A implementación realizouse tendo en conta diversas estratexias de optimización e sen poñer en risco os bos resultados do método a nivel de precisión do rexistro. Destaca o uso da biblioteca cuFFT para acelerar o cálculo das numerosas FFTs empregadas polo método e a utilización eficiente dos diferentes espazos de memoria, como por exemplo, o uso da memoria de texturas que permite realizar a interpolación de datos na propia instrución de acceso a memoria. Os resultados son avaliados en termos de tempos de execución e de ocupancia da GPU.

- No Capítulo 5 proponse o primeiro método que implementa o esquema de rexistro adaptable a calquera detector e descritor de características. Concretamente, neste capí-

tulo preséntase un método para o rexistro de dúas imaxes hiperespectrais baseado en características chamado Hyperspectral KAZE (HSI–KAZE). A proposta adapta o algoritmo KAZE e o algoritmo Accelerated-KAZE (A-KAZE) para o seu uso en imaxes hiperespectrais. O método basease na detección de puntos de interese en diferentes bandas espectrais. Este emprega unha difusión non lineal para construír a escala-espazo aplicando un suavizado controlado e M-SURF como descritor espacial dos puntos de interese. Este descritor é enriquecido cunha parte espectral. Como parte do esquema de rexistro proponse un método de selección de bandas e un método para estimar a transformación xeométrica a partir dos puntos de interese previamente emparellados. Ambas técnicas son comparadas con outras dispoñibles na literatura demostrándose ademais que as nosas propostas acadan mellores resultados. HSI–KAZE é avaliado empregando imaxes baixo distintas condicións e tipos, e comparado con outros métodos da literatura.

- No Capítulo 6 preséntase unha implementación eficiente en GPU de HSI–KAZE. No capítulo descríbense as diferentes técnicas e estratexias de optimización empregadas para acadar tempos que posibilitan a execución do método en tempo real. Destaca, entre todas elas, o uso de reducións a nivel de *warp* e de bloque empregando as novas primitivas e instrucións atómicas a nivel de *warp* e de memoria compartida que fan o intercambio de información máis eficiente ao reducir o número de colisións en memoria global. A implementación en GPU acada os mesmos bos resultados que a implementación en CPU. Esta posibilita o uso deste custoso método en CPU en aplicacións onde o tempo de execución é crucial. O método é ademais comparado con unha implementación multifío en CPU usando OpenMP.

- No Capítulo 7 preséntase Hyperspectral SURF (HSI–SURF), un método de rexistro de imaxes hiperespectrais baseado en *Speeded Up Robust Features (SURF)*. Trátase do segundo método que implementa o esquema de rexistro proposto e no que destaca o método de selección de bandas, a explotación de distintas bandas espectrais, a incorporación da información espectral no descritor e o método de estimación da transformación xeométrica que permite aliñar as imaxes. O método incorpora a SURF a información espectral coa fin de acadar aliñamentos de imaxes con grandes escalas pero sen perder a fin pola que o método orixinal foi creado, a eficiencia computacional. Neste traballo vólvese a probar a eficacia do método de selección de bandas proposto para o rexistro de imaxes hiperespectrais e que está integrado no propio esquema.

- No Capítulo 8 preséntase Hyperspectral MSER (HSI–MSER), un método de rexistro de características de imaxes hiperespectrais. É o terceiro método que implementa o esquema de rexistro proposto. En contraste cos outros métodos basados en características presentados, HSI–MSER procura por rexións comúns nas imaxes a rexistrar en vez de puntos de interese. O método está baseado no método *Maximally Stable Extremal Regions (MSER)* para a detección de ditas rexións e nunha adaptación do descritor de SIFT para a descrición espacial de rexións no canto de puntos de interese. HSI–MSER incorpora todas as melloras e propostas do propio esquema acadando bos resultados tanto en termos de precisión como en tempos de computación.

- Finalmente, no Capítulo 9 proponse un esquema multi-GPU de rexistro de imaxes multiespectrais de dous niveis. No primeiro nivel rexístranse as bandas de cada imaxe multiespectral. No segundo nivel, unha vez aliñadas as bandas, rexístranse as diferentes imaxes multiespectrais. O rexistro é realizado de maneira independente por dous métodos, HYFM e HSI–KAZE. Desta maneira ambos métodos son comparados en termos de precisión nun problema de rexistro distinto ao de anteriores experimentos. Diferentes conxuntos de datos multiespectrais, pero que forman parte da mesma escena a rexistrar, son distribuídos a diferentes GPUs empregando OpenMP. Cada GPU realiza en paralelo o rexistro dun conxunto de datos diferente. Deste xeito realízase tamén paralelismo a dous niveis, a nivel de GPU, pois diferentes conxuntos de datos son aliñados en paralelo, e a nivel interno de cada GPU grazas ás implementacións eficientes realizadas. Esta implementación multi-GPU é comparada con outra implementación multi-CPU en OpenMP.

En vista dos resultados obtidos, podemos concluír que se acadaron os obxectivos desta tese. Confirmáronse ambas hipóteses, a información espectral permite aliñar imaxes multi e hiperespectrais con grandes e descoñecidos escalados e rotacións, así como o problema de rexistro automático deste tipo de imaxes pode realizarse en tempo real grazas ao desenvolvemento de algoritmos eficientes e paralelos para GPU.

Resumen

El gran avance en el desarrollo de sensores permite que hoy podamos obtener imágenes multiespectrales e hiperespectrales de teledetección a menores costes que antaño. La característica más significativa de estas imágenes es el rango de longitud de onda que abarcan [13]. Desde el visible, pasando por el infrarrojo cercano y hasta el infrarrojo medio del espectro electromagnético, es decir, desde 0.4 a 2.5 μm [34]. Las imágenes multi e hiperespectrales están formadas por muchos canales espectrales continuos, también llamados bandas. Una imagen con 10 o menos bandas recibe el nombre de imagen multiespectral, mientras que una imagen con 10 o más bandas, pero habitualmente más de 100, recibe el nombre de imagen hiperespectral. Esto significa que cada píxel es un vector donde cada componente se corresponde con una determinada longitud de onda [49]. Gracias a la alta resolución espectral de estas imágenes podemos distinguir con mayor precisión objetos, especies, usos del suelo, entre otros.

Cientos de imágenes hiperespectrales son obtenidas todos los días gracias al creciente número de satélites y aeronaves con sensores espectrales [20, 25, 61]. Esto nos permite obtener imágenes de la misma zona de la Tierra tomadas en distintos instantes de tiempo (minutos, horas, meses o incluso años) y desde diferentes perspectivas. Esta serie de imágenes es empleada en tareas donde es esencial comparar, estudiar o buscar las diferencias entre imágenes. Algunos ejemplos de estas tareas o aplicaciones son la detección automática de cambios [32, 59, 77, 79], el control medioambiental [40, 120, 138], la fusión de imágenes [7, 14, 63, 87] o la creación de imágenes con superresolución [4, 68, 69]. Antes de llevar a cabo cualquier procesado asociado a dichas tareas, las imágenes tienen que ser registradas.

El registro, alineamiento o superposición de imágenes consiste en alinear imágenes de la misma zona o área pero que fueron obtenidas en distintos momentos temporales, desde diferentes perspectivas, y posiblemente bajo diferentes condiciones lumínicas. Si nos limitamos al registro de dos imágenes, lo que se pretende es registrar o alinear una de ellas, llamada

imagen de referencia, con respeto de la otra, llamada imagen objetivo. Esto exige localizar y emparejar en ambas imágenes los mismos objetos, estructuras, regiones, etc. El registro no es un proceso sencillo pues habitualmente las imágenes presentan cambios debido al paso del tiempo desde que estas fueron tomadas, pero también presentan cambios de iluminación, de resolución espacial y espectral, etc., y puede que no cubran exactamente la misma región.

A todo esto hay que sumarle otras dificultades asociadas a las imágenes de teledetección. Estas imágenes presentan ciertas características que hacen del registro un tarea difícil y compleja en comparación con otros tipos de imágenes, por ejemplo, las imágenes médicas [71]. Los algoritmos de registro tienen que lidiar con la amplia variedad de sensores y con las distintas condiciones en la que las imágenes pueden ser obtenidas; con las condiciones atmosféricas y meteorológicas; con la estación del año en la que nos encontremos ya que pueden hacer que elementos naturales como árboles, lagunas y ríos sean puntos de referencia nada o poco fiables para el registro; con las diferentes condiciones de luz debido a efectos del paso del tiempo tal como el cambio del ángulo del Sol durante el año; y la grande cantidad de datos. Todo esto hace del registro de imágenes de teledetección terrestre una difícil tarea.

El registro de imágenes puede hacerse a la mano. Habitualmente se realiza seleccionando puntos característicos (puntos de control) comunes en todas las imágenes a registrar. El registro manual no es viable cuando hablamos de imágenes de teledetección y más específicamente cuando hablamos de imágenes tomadas por satélites, los cuales, capturan cientos de imágenes al día. El proceso exigiría un alto coste temporal y de personal para llevarlo a cabo. Además, el registro manual no es sencillo de realizar pues muchas veces es difícil encontrar puntos de referencia debido a la propia resolución de las imágenes o porque se tratan de imágenes de urbes o zonas agrícolas con mucha monotonía. Nace así el registro automático como un área de investigación en el procesamiento de imágenes.

Los métodos de registro automático pueden clasificarse en dos categorías [149]:

- **Métodos basados en área**. Estos métodos se fundamentan en el cómputo de diferencias entre píxeles [71]. Los métodos basados en área explotan directamente los datos de las imagenes sin ningún análisis estructural, es decir, no buscan correspondencias entre las imágenes buscando puntos o regiones comunes. Estes métodos pueden a su vez dividirse en tres categorías: métodos basados en información mutua [15, 65], métodos basados en la correlación [9, 62] y métodos basados en la transformada de Fourier [16, 122]. Los métodos basados en información mutua precisan que la intensidad entre las imágenes sea estable, lo cual no es una característica muy común en las imágenes

de teledetección terrestre. Por esta razón, estos métodos son habitualmente empleados en imagen médica donde las condiciones para la obtención de las imágenes está muy controlada al ser adquiridas en laboratorios o hospitales. Los métodos de la segunda categoría, los métodos basados en la correlación, se basan en optimizar una medida de semejanza desplazando regiones de la imagen de referencia a la imagen objetivo. La medida de semejanza se computa para pares de ventanas de ambas imágenes lo que hace que este tipo de métodos sean computacionalmente caros.

Por lo contrario, los métodos basados en la transformada de Fourier son una alternativa muy eficiente en términos computacionales para calcular la correlación entre imágenes. Estos métodos emplean el teorema del desplazamiento para la transformada de Fourier el cual indica que un desplazamiento circular en el dominio espacial equivale a una diferencia de fase en el dominio de la frecuencia. Este tipo de métodos son una buena elección cuando se trata de registrar imágenes que fueron obtenidas bajo diferentes condiciones [71, 149]. Estas características hacen de los métodos basados en la transformada de Fourier en una opción interesante para el registro de imágenes hiperespectrales de teledetección terrestre. En general, los métodos basados en área son computacionalmente muy eficientes pero su precisión puede verse afectada por oclusiones, cambios temporales, ruido, u otros efectos que pueden producir cambios en la intensidad de los píxeles.

- **Métodos basados en características**. Este tipo de métodos buscan por características distintivas o de interés en las imágenes a registrar. Estas características pueden ser puntos, contornos, regiones o líneas que cumplen con una serie de requisitos: ser invariantes a transformaciones geométricas, fáciles de localizar en términos de precisión y ser insensitivas a degradaciones. Estos métodos se basan en detectar las mismas características en las imágenes a registrar. De esta manera, conociendo un cierto número de características comunes, se puede calcular una transformación geométrica que registre una imagen con respeto de la otra. Esta búsqueda de características comunes a alto nivel hace que este tipo de métodos se vean menos afectados por cambios de iluminación, de intensidad por la presencia de ruido o por tratarse de un registro de imágenes capturadas por diferentes sensores [53]. Además, las imágenes de teledetección suelen contener muchas líneas, esquinas, regiones, etc., susceptibles de ser detectadas por este tipo de métodos. Estas razones convierten a los métodos de registro basados en características en los más empleados en cuanto a imágenes de teledetección.

> Por el contrario, estos métodos requieren de la construcción de una escala-espacio para detectar características con los requisitos descritos, así como para describir cada una de ellas mediante un vector único pero que no se vea afectado por transformaciones geométricas, cambios de luz o ruido. Esto se traduce en un alto coste computacional.

En la categoría de métodos basados en área destaca el *Fourier-Mellin invariant symmetric phase-only matched filtering (FMI-SPOMF)* [16] mientras que en los métodos basados en características destaca *Scale-Invariant Feature Transform (SIFT)* [80]. Estos algoritmos fueron desarrollados para procesar imágenes en escala de grises o imágenes *RGB* (del inglés, *Red Blue Green*) y no para lidiar con la información espectral disponible en las imágenes multi e hiperespectrales. La información espectral de estas imágenes nos brindan una información extra que puede ser empleada para mejorar la precisión del registro. Hay características que están sólo presentes en alguna de las bandas de las imágenes y que pueden ser empleadas a la hora de computar la transformación geométrica que buscamos para alinear las imágenes. Además, la información espectral, junto con la información espacial, puede ser utilizada para esa distinción inequívoca de si un punto o región de una imagen es o no es el equivalente en otra imagen. El uso de esta información parece más que justificado pero tiene un defecto, el alto coste computacional de su procesamiento. El registro, al igual que cualquiera otra tarea habitual de procesado de imágenes multi e hiperespectrales (detección de cambios, clasificación, detección de objetos, etc.), requiere de un alto coste temporal. Algunas propuestas de la literatura [58, 74, 85, 132, 143] resuelven este defecto procesando las imágenes multi e hiperspectrales como imágenes de una sola banda perdiendo de esta forma todo el potencial de este tipo de imágenes y renunciando así a emplear estos datos para realizar registros más precisos o capaces de alinear imágenes con mayor diferencia de escala, rotación o traslación.

En otros trabajos [24, 88] se propusieron métodos de registro teniendo en cuenta la información espectral, pero ninguno de ellos resuelve el problema de registro eficiente de grandes conjuntos de imágenes hiperespectrales o su realización en tiempo real. El tiempo de cómputo es crucial en muchos casos de uso, por ejemplo, en el control de desastres medioambientales. Además, el número de imágenes hiperespectrales en las bases de datos de las agencias espaciales no deja de crecer cada día que pasa, evidenciando la necesidad de métodos eficientes para procesar esa gran cantidad de datos. Es también una realidad que, gracias al abaratamiento de los precios de los espectrómetros, su uso se está extendiendo a ámbitos más humildes como el caso de los vehículos aéreos no tripulados (en inglés, *unmanned aerial vehicles (UAVs)*). Todo esto hace indispensable el desarrollo de algoritmos de registro

automáticos que sean eficientes tanto aprovechando y explotando la información espectral para mejorar la calidad del registro como desde el punto de vista computacional.

Es por todo esto que se definen como hipótesis de esta tesis las siguientes:

H1. *El registro automático de imágenes multi e hiperespectrales con grandes diferencias a nivel escala y orientación, además de cambios, por ejemplo, fruto del paso del tiempo, es posible gracias a la explotación eficiente de la información espectral.*

H2. *El registro automático de imágenes multi e hiperespectrales puede ser realizado en tiempo real gracias al desarrollo de algoritmos paralelos en hardware de bajo coste.*

El proceso de demostración de esas hipótesis dan lugar el objetivo principal de esta tesis. Este es el desarrollo de algoritmos eficientes de registro automático de imágenes multi e hiperespectrales de teledetección de cobertura terrestre. Para eso, los algoritmos desarrollados emplean la gran cantidad de información que estas imágenes contienen, tanto de tipo espectral cómo espacial, consiguiendo alinear imágenes con escalas, rotaciones y traslaciones superiores con respecto a trabajos publicados en la literatura.

En esta tesis doctoral se propone un método basado en la transformada de Fourier así como un esquema de registro adaptable a cualquier detector y descriptor de características. Este esquema fue la base de desarrollo de tres algoritmos de registro de imágenes hiperespectrales basados en detección de características. También se desarrollaron métodos específicos para la selección de bandas y el cómputo de transformación geométrica a partir de la extracción de características. Ambos forman parte del mencionado esquema de registro.

Además, se han realizado implementaciones eficientes en unidades de procesamiento gráfico (en inglés, *Graphics Processing Unit (GPU)*) que nos permiten registrar imágenes en tiempo real así como grandes conjuntos de datos empleando *hardware* de bajo coste, es decir, sin precisar de una infraestructura de alto rendimiento como son los supercomputadores. Se presenta también una implementación multi-GPU para la realización de un registro multinivel en paralelo de diferentes conjuntos de imágenes multiespectrales.

Para realizar la experimentación, se construyó un conjunto de datos de prueba que está disponible para toda la comunidad científica [115]. Este conjunto está formado por pares de imágenes tomadas por los sensores ROSIS y AVIRIS en distintas fechas, desde distintas perspectivas y que presentan cambios fruto del paso del tiempo y de iluminación así como distorsiones. Los algoritmos y métodos desarrollados en esta tesis fueron comparados con métodos del estado del arte en términos de precisión y rendimiento obteniendo buenos resul-

tados. Se consiguió registrar imágenes con escalados de hasta $24\times$ en un tiempo de ejecución de entre 1 y 4 segundos de promedio.

Esta tesis doctoral ha dado lugar a diversas publicaciones en revistas y congresos internacionales. Las más relevantes son recogidas en los siguientes capítulos de esta tesis:

- En el Capítulo 3 se propone Hyperspectral Fourier–Mellin (HYFM), un método para el registro de dos imágenes hiperespectrales basado en la transformada de Fourier, los métodos de área más eficientes. El método emplea la información contenida en las diferentes bandas a lo largo de sus seis etapas. La transformada rápida de Fourier (en inglés, *fast Fourier transform (FFT)*) es utilizada para el cómputo eficiente de la correlación de las dos imágenes. Se hace uso de la transformada fraccional de Fourier (en inglés, *fractional Fourier transform (FRFT)*) y de la transformada fraccional multicapa de Fourier (en inglés, *multi-layer fractional Fourier transform (MLFFT)*) para obtener registros más precisos pues se reducen los errores de interpolación gracias al cómputo de una cuadrícula multinivel. El método computa la correlación en ocho pares de componentes principales calculadas en la primera etapa empleando la técnica de análisis de componentes principales (en inglés, *Principal Component Analysis (PCA)*). Esta información es luego combinada en la quinta etapa para así tener en cuenta toda la información espectral a la hora de computar la transformación geométrica que registre las imágenes. El método es comparado con otros métodos en la literatura así como otros métodos de extracción de características distintos del análisis de componentes principales.

- En el Capítulo 4 se presenta una implementación eficiente en GPU del método HYFM para el registro de imágenes hiperspectrales. El método es ejecutado totalmente en GPU explotando así su arquitectura de cómputo y evitando las transferencias de datos innecesarios entre la CPU y la GPU que lastrarían el rendimiento. La implementación se realizó teniendo en cuenta diversas estrategias de optimización y sin poner en riesgo los buenos resultados del método a nivel de precisión del registro. Destaca el uso de la biblioteca cuFFT para acelerar el cálculo de las numerosas FFTs empleadas por el método y la utilización eficiente de los diferentes espacios de memoria, como por ejemplo, el uso de la memoria de texturas que permite realizar la interpolación de datos en la propia instrucción de acceso a memoria. Los resultados son evaluados en términos de tiempos de ejecución y de ocupancia de la GPU.

- En el Capítulo 5 se proponen el primer método que implementa el esquema de registro adaptable a cualquier detector y descriptor de características. Concretamente, en este capítulo se presenta un método para el registro de dos imágenes hiperespectrales basado en características llamado Hyperspectral KAZE (HSI–KAZE). La propuesta adapta el algoritmo KAZE y el algoritmo Accelerated-KAZE (A-KAZE) para su uso en imágenes hiperespectrales. El método se basa en la detección de puntos de interés en diferentes bandas espectrales. Este emplea una difusión no lineal para construir la escala-espacio aplicando un suavizado controlado y M-SURF como descriptor espacial de los puntos de interés. Este descriptor es enriquecido con una parte espectral. Como parte del esquema de registro se proponen un método de selección de bandas y un método para estimar la transformación geométrica a partir de los puntos de interés previamente emparejados. Ambas técnicas son comparadas con otras disponibles en la literatura demostrándose además que nuestras propuestas alcanzan mejores resultados. HSI–KAZE es evaluado empleando imágenes bajo distintas condiciones y tipos, y comparado con otros métodos de la literatura.

- En el Capítulo 6 se presenta una implementación eficiente en GPU de HSI–KAZE. En el capítulo se describen las diferentes técnicas y estrategias de optimización empleadas para alcanzar tiempos que posibilitan la ejecución del método en tiempo real. Destaca, entre todas ellas, el uso de reducciones a nivel de *warp* y de bloque empleando las nuevas primitivas e instrucciones atómicas a nivel de *warp* y de memoria compartida que hacen el intercambio de información más eficiente al reducir el número de colisiones en memoria global. La implementación en GPU alcanza los mismos buenos resultados que la implementación en CPU. Esta posibilita el uso de este costoso método en CPU en aplicaciones donde el tiempo de ejecución es crucial. El método es además comparado con una implementación multifío en CPU usando OpenMP.

- En el Capítulo 7 se presenta Hyperspectral SURF (HSI–SURF), un método de registro de imágenes hiperespectrales basado en *Speeded Up Robust Features (SURF)*. Se trata del segundo método que implementa el esquema de registro propuesto y en el que destaca el método de selección de bandas, la explotación de distintas bandas espectrales, la incorporación de la información espectral en el descriptor y el método de estimación de la transformación geométrica que permite alinear las imágenes. El método incorpora a SURF la información espectral con el fin de alcanzar alineamientos de imágenes con grandes escalas pero sin perder el fin por la el cual el método original fue creado, la

eficiencia computacional. En este trabajo se vuelve a probar la eficacia del método de selección de bandas propuesto para el registro de imágenes hiperespectrales y que está integrado en el propio esquema.

- En el Capítulo 8 se presenta Hyperspectral MSER (HSI–MSER), un método de registro de características de imágenes hiperespectrales. Es el tercer método que implementa el esquema de registro propuesto. En contraste con los otros métodos basados en características presentados, HSI–MSER busca por regiones comunes en las imágenes a registrar en vez de puntos de interés. El método está basado en el método *Maximally Stable Extremal Regions (MSER)* para la detección de dichas regiones y en una adaptación del descriptor de SIFT para la descripción espacial de regiones en vez de puntos de interés. HSI–MSER incorpora todas las mejoras y propuestas del propio esquema alcanzando buenos resultados tanto en términos de precisión como en tiempos de computación.

- Finalmente, en el Capítulo 9 se proponen un esquema multi-GPU de registro de imágenes multiespectrales de dos niveles. En el primero nivel se registran las bandas de cada imagen multiespectral. En el segundo nivel, una vez alineadas las bandas, se registran las diferentes imágenes multiespectrales. El registro es realizado de manera independiente por dos métodos, HYFM y HSI–KAZE. De esta forma ambos métodos son comparados en términos de precisión en un problema de registro distinto al de anteriores experimentos. Diferentes conjuntos de datos multiespectrales, pero que forman parte de la misma escena a registrar, son distribuidos a diferentes GPUs empleando OpenMP. Cada GPU realiza en paralelo el registro de un conjunto de datos diferente. De esta manera se realiza también paralelismo a dos niveles, a nivel de GPU, pues diferentes conjuntos de datos son registrados en paralelo, y a nivel interno de cada GPU gracias a las implementaciones eficientes realizadas. Esta implementación multi-GPU es comparada con otra implementación multi-CPU en OpenMP.

En vista de los resultados obtenidos, podemos concluir que se alcanzaron los objetivos de esta tesis. Se confirmaron ambas hipótesis, la información espectral permite alinear imágenes multi e hiperespectrales con grandes y desconocidos escalados y rotaciones, así como el problema de registro automático de este tipo de imágenes puede realizarse en tiempo real gracias al desarrollo de algoritmos eficientes y paralelos para GPU.

Summary

The breakthroughs in sensor development allow obtaining multispectral and hyperspectral remote sensing images at lower costs than in the past. The most significant characteristic of these images is the wavelength range they cover [13]: from the visible through the near-infrared to the mid-infrared of the electromagnetic spectrum, i.e. from 0.4 to $2.5\,\mu\mathrm{m}$ [34]. Multi and hyperspectral images consist of many spectral channels, also called bands. An image with 10 or fewer bands is called a multispectral image, while an image with 10 or more bands, but usually more than 100, is called a hyperspectral image. This means that each pixel can be represented as a vector where each component corresponds to the intensity of a specific wavelength [49]. Owing to the high spectral resolution of these images, we can more accurately distinguish objects, species, land use, among others.

Hundreds of hyperspectral images are obtained every day owing to the growing number of satellites and aircraft with spectral sensors [20, 25, 61]. This allows us to obtain images of the same area of the Earth taken at different times (minutes, hours, months or even years) and from different perspectives. These series of remote-sensing images are used in tasks where it is essential to compare, study, or search for differences between images. Examples of such tasks or applications are automatic change detection [32, 59, 77, 79], environmental monitoring [40, 120, 138], image fusion [7, 14, 63, 87], or super-resolution image creation [4, 68, 69]. Before any image processing associated with these tasks can be performed, the images have to be registered.

Image registration consists in aligning images of the same area which were obtained at different times, from different perspectives, and possibly under different light conditions. Considering the problem of the registration of two images, the aim is to register one of them, called the reference image, with respect to the other, called the target image. This requires locating and matching the same objects, structures, regions, etc., in both images. Registration

is not a trivial process because the images usually present changes due to the passage of time since each one was taken, as well as changes in illumination, spatial and spectral resolution, etc., and may not cover exactly the same region.

Moreover, there are other difficulties related to remote sensing images. Remote sensing images have some characteristics that make their registration challenging as compared to other types of images, e.g. medical images [71]. Registration algorithms have to deal with the wide variety of sensors and data acquisition conditions; atmospheric and meteorological conditions; the season of the year, which can make natural elements such as trees, lakes, or rivers unreliable references for registration; different light conditions due to weather effects such as the change of the Sun's angle during the year; and the large amount of data. All this makes the registration of remote sensing images a difficult task.

Image registration can be done by hand. It is usually done by selecting distinctive points (control points) present in all images. Manual registration is not feasible when dealing with remote sensing images, and particularly when dealing with images taken by satellites, which capture hundreds of images per day. The process would require a high cost in terms of time and staff to carry it out. Furthermore, it is often difficult to find reference points due to the resolution of the images or because they are images of very monotonous urban or rural areas. Thus, automatic registration was born as an area of research in image processing.

Automatic image registration methods can be classified into two categories [149]:

- **Area-based methods**. These methods are based on the computation of the differences between pixels [71]. They directly exploit the image intensities without any structural analysis, i.e. they do not look for correspondences between images by looking for common points or regions. Area–based methods can be subdivided into three categories: mutual information (MI) methods [15, 65], correlation-based approaches [9, 62], and Fourier transform methods [16, 122]. MI-based methods require that the intensities between images be stable, which is not a very common feature in remote sensing images. For this reason, these methods are generally used in medical images, in which the imaging conditions are tightly controlled, as they are acquired in laboratories or hospitals. Methods in the second category, correlation-based methods, are based on optimising a similarity measure by translating regions of the reference image to the target image. The similarity measure is computed for window pairs for both the images, which makes correlation-based methods very computationally expensive.

  In contrast, Fourier-based methods are a computationally efficient alternative for calcu-

lating the correlation between images. These methods rely on the Fourier shift theorem, which states that a circular shift in the spatial domain is equivalent to a phase ratio in the frequency domain. Fourier-based methods are a good choice when images have been taken under different conditions [71, 149]. These characteristics make Fourier-based methods an interesting option for the registration of hyperspectral remote sensing images. In general, area-based methods are computationally very efficient but their accuracy can be affected by occlusions, temporal changes, noise, or other effects that can produce changes in pixel intensities.

- **Feature-based methods**. These methods look for distinctive features in the images. These features can be points, contours, regions or lines that must have certain characteristics: invariant to geometric transformations, good localisation accuracy, and insensitivity to image degradation effects. Feature-based methods rely on the extraction of the same features in the images to be registered. On the basis of the knowledge of a number of corresponding features, a geometric transformation that aligns one image with respect to the other can be calculated. The search for common features at a high level makes this type of method more resilient to changes in illumination, noise, and multisensor registration [53]. Moreover, remote sensing images often contain many lines, corners, regions, etc. that can be detected by feature-based methods. These reasons make feature-based registration algorithms the most widely used methods for the registration of remote sensing images.

  In contrast, they require the construction of a scale-space to detect features with the described characteristics, as well as to describe each feature by a unique vector that is not affected by geometric transformations, light changes, or noise. This results in a high computational cost.

Within the category of area-based methods, the Fourier-Mellin invariant symmetric phase-only matched filtering (FMI-SPOMF) method [16] stands out, while in feature-based methods, the Scale-Invariant Feature Transform (SIFT) method [80] stands out. These algorithms were developed to process greyscale or RGB images and not to deal with the spectral information available in multi and hyperspectral images. The spectral information in these images provides additional information that can be used to improve the accuracy of the registration. There are features that are only present in some of the bands that can be used to compute the geometric transformation that aligns the images. Moreover, the spectral information, with the

spatial information, can be used to check whether or not a point or region in one image is the equivalent one in another image. The use of this information seems more than justified, but it has a drawback: the high computational cost of processing it. Registration, like any other common multi and hyperspectral image processing task (change detection, classification, object detection, etc.), requires a high time cost. Some proposals in the literature [58, 74, 85, 132, 143] solve this issue by processing only one band of the multi and hyperspectral images, thus renouncing the use of the spectral information for more accurate registration or for aligning images with larger differences in scale, rotation or translation.

Other works [24, 88] proposed registration methods taking into account the spectral information, but none of them solve the problem of efficient registration of large hyperspectral image datasets or their real-time application. The computational time is crucial in many use cases, for example, in environmental disaster monitoring. Furthermore, the number of hyperspectral images in space agency databases is growing every day, becoming evident the need for efficient methods to process this large amount of data. It is also a reality that owing to the lower prices of spectrometers, their use is extending to humble settings, for example, unmanned aerial vehicles (UAVs). All this makes it essential to develop automatic registration algorithms that are efficient both in terms of exploiting spectral information to improve the registration accuracy and from the computational point of view.

For all these reasons, the hypotheses of this thesis are defined as follows:

H1. *Exploiting the spectral information available in the different bands allows the alignment of multispectral and hyperspectral images with large and unknown scale factors and rotations.*

H2. *Registration of multispectral and hyperspectral images may be carried out in real-time by developing efficient and parallel algorithms for commodity hardware.*

The process of testing these hypotheses leads to the main objective of this thesis. The main goal is the development of efficient algorithms for the automatic registration of multi and hyperspectral remote sensing images. To this end, the developed algorithms use the large amount of information that hyperspectral images contain, both spectral and spatial, achieving the alignment of images with scales, rotations, and translations that are higher than those published in the literature.

In this thesis, a Fourier-based method and a feature-based registration pipeline adaptable to any feature detector and descriptors are proposed. The pipeline was the basis for the devel-

opment of three feature-based registration algorithms. Specific methods for band selection, and for choosing the matched features and calculating the final geometric transformation were also developed. Both are part of the above-mentioned registration pipeline.

Moreover, efficient implementations have been carried out on Graphics Processing Units (GPUs) using CUDA. It allows us to register images in real-time, as well as large datasets, using low-cost hardware, i.e. without the need for high-performance infrastructures such as supercomputers. A multi-GPU implementation for parallel multilevel registration of different datasets of multispectral images is also presented.

To carry out the experimentation, a dataset of hyperspectral images were created to test registration algorithms. It is available online for the scientific community [115] and consists of pairs of images taken by the ROSIS and AVIRIS sensors on different dates, from different perspectives, and that present changes due to the passage of time, illumination, and distortions. The algorithms and methods developed in this thesis were compared to state-of-the-art methods in terms of accuracy and performance, obtaining good results. It was possible to register images with scale factors of up to $24\times$ on an average execution time of 1 to 4 seconds.

This doctoral thesis has resulted in several publications in international journals and conferences. The most relevant ones are collected in the following chapters of this thesis:

- Chapter 3 proposes the Hyperspectral Fourier–Mellin (HYFM) method, a method for the registration of two hyperspectral images based on the Fourier transform, the most efficient area-based methods. It uses the information contained in the different bands throughout its six stages. The fast Fourier transform (FFT) is used for the efficient computation of the correlation of the two images. The fractional Fourier transform (FRFT) and the multi-layer fractional Fourier transform (MLFFT) are used to obtain more accurate registrations, as interpolation errors are reduced due to the computation of a multilevel grid. The method computes the correlation in eight pairs of principal components calculated in the first stage using Principal Component Analysis (PCA). This information is then combined in the fifth stage. Thus, all the spectral information is taken into account when the geometric transformation, which aligns the images, is computed. The method is compared to others published in the literature. Additionally, feature extraction methods other than PCA are evaluated.

- Chapter 4 presents an efficient GPU implementation of the HYFM method for hyperspectral image registration. The method is fully executed on GPU, taking advantage of its architecture and avoiding unnecessary data transfers between the CPU and GPU that

would imply a performance penalty. The implementation was carried out taking into account several optimisation strategies and without compromising the good results of the method in terms of registration accuracy. The use of the cuFFT library to speed up the calculation of the numerous FFTs used by the method and the efficient use of the different memory spaces stand out. An example of the last point is the use of the texture memory which allows the interpolation of data in the memory access instruction itself. The results are evaluated in terms of execution times and GPU occupancy.

- Chapter 5 presents the first method that implements the proposed feature-based registration pipeline adaptable to any feature detector and descriptor. In particular, this chapter presents a feature-based method for the registration of two hyperspectral images called Hyperspectral KAZE (HSI–KAZE). The proposal adapts the KAZE and the Accelerated-KAZE (A-KAZE) methods for use in hyperspectral images. It is based on the detection of keypoints in different spectral bands. It employs a non-linear diffusion to build the scale-space by applying controlled smoothing and M-SURF as the spatial descriptor of the keypoints. This descriptor is enriched with a spectral part. As part of the registration pipeline, a band selection method and a method to estimate the geometric transformation from the matched keypoints are proposed. Both techniques are compared to others available in the literature and it is shown that our proposals achieve better results. HSI–KAZE is evaluated using images under different conditions and types, and compared to other methods in the literature.

- Chapter 6 proposes an efficient GPU implementation of HSI–KAZE. The chapter describes the different optimisation techniques and strategies used to achieve real-time execution. Among all of them, the one that stands out the most is the use of warp and block-level reductions using the new primitives and atomic instructions at the warp and shared memory levels that make the exchange of information more efficient by reducing the number of collisions in global memory. The GPU implementation achieves the same good results as the CPU implementation. The GPU implementation makes it possible to use this CPU-intensive method in time-critical applications. Furthermore, the method is compared to a multicore OpenMP implementation.

- Chapter 7 presents Hyperspectral SURF (HSI–SURF), a hyperspectral image registration method based on Speeded Up Robust Features (SURF). It is the second method that implements the proposed registration pipeline. It stands out for the band selection

method, the exploitation of different spectral bands, the incorporation of the spectral information in the descriptor, and the method for estimating the geometric transformation. The method adapts SURF to incorporate the spectral information for achieving image alignments with large scales but without losing the purpose for which the original method was created, computational efficiency. In this work, the effectiveness of the proposed band selection method for hyperspectral image registration, which is integrated into the pipeline itself, is also evaluated.

- Chapter 8 proposes Hyperspectral MSER (HSI–MSER), a feature-based registration method for hyperspectral images. It is the third method that implements the proposed registration pipeline. In contrast to the other feature-based methods developed, HSI–MSER searches for common regions in the images instead of keypoints. The method is based on the Maximally Stable Extremal Regions (MSER) method for the detection of regions and on an adaptation of the SIFT descriptor for the spatial description of regions instead of keypoints. HSI–MSER incorporates all the improvements and proposals of the pipeline and achieves good results both in terms of accuracy and computational time.

- Finally, in Chapter 9 a two-level multi-GPU system for the registration of different datasets of multispectral images is proposed. In the first level, the bands of each multispectral image are registered. In the second level, once the bands have been aligned, the different multispectral images are registered. The registration is performed independently by two methods, HYFM and HSI–KAZE. Thus, both the methods are compared in terms of accuracy in a different registration problem. Different multispectral datasets, but which are part of the same scene, are distributed to different GPUs using OpenMP. Each GPU registers a different dataset in parallel. Thus, parallelism is also carried out at two levels, at the GPU level, as different datasets are registered in parallel, and at the internal level of each GPU owing to the efficient implementations developed. The multi-GPU implementation is compared to another multi-CPU implementation in OpenMP.

In view of the results obtained, we can conclude that the objectives of this thesis were achieved. Both hypotheses were confirmed, the spectral information allows registering multi and hyperspectral images with large and unknown scales and rotations. Moreover, the problem of automatic registration of this type of images can be performed in real-time owing to the development of efficient and parallel algorithms for GPU.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Contents

Acronyms

A-KAZE Accelerated-KAZE.

API Application Programming Interface.

AVIRIS Airborne Visible/Infrared Imaging Spectrometer.

BLAS Basic Linear Algebra Subroutines.

CPU Central Processing Unit.

CUDA Compute Unified Device Architecture.

EBS Entropy-Based Band Selection.

FED Fast Explicit Diffusion.

FFT fast Fourier transform.

FRFT fractional Fourier transform.

GPC Graphics Processing Cluster.

GPU Graphics Processing Unit.

HSI hyperspectral image.

HSI–KAZE Hyperspectral KAZE.

HSI–MSER Hyperspectral MSER.

HSI–SURF Hyperspectral SURF.

HYFM Hyperspectral Fourier–Mellin.

M-SURF Modified-SURF.

MLFFT multi-layer fractional Fourier transform.

MSER Maximally Stable Extremal Regions.

OpenMP Open Multi-Processing.

PC principal component.

PCA Principal Component Analysis.

RANSAC Random Sample Consensus.

RCMG Robust Colour Morphological Gradient.

RMSE Root-Mean-Square Error.

ROSIS Reflective Optics System Imaging Spectrometer.

SIFT Scale-Invariant Feature Transform.

SIMT Single Instruction, Multiple Thread.

SM Streaming Multiprocessor.

SPMD Single Program, Multiple Data.

SURF Speeded Up Robust Features.

TPC Texture Processing Cluster.

WaluMI Ward's Linkage Strategy Using Mutual Information.

Chapter 1

Introduction

This thesis is organised into ten chapters. Chapter 1 introduces the fundamental concepts involved in this work, its motivation, the hypothesis and objectives, as well as the hardware and methodology followed to carry it out. In particular, Section 1.1 presents the main characteristics of multispectral and hyperspectral remote sensing images and their acquisition, the fundamentals of image registration, its applications, the different types of algorithms that exist, as well as a review of the state-of-the-art methods. The hypothesis and the main objectives of this doctoral thesis are detailed in Section 1.2. Section 1.3 presents the basics of parallel programming using Open Multi-Processing (OpenMP) and Compute Unified Device Architecture (CUDA), and the hardware and the software used. This section also describes the multispectral and hyperspectral images used to evaluate the proposed methods as well as the performance measures used for this purpose. Finally, the publications resulting from this thesis are listed in Section 1.4.

Chapter 2 presents the main contributions of this thesis. A feature-based method and a feature-based registration pipeline, which is used as the basis in three different methods, are presented. The proposed algorithms and their results are discussed in comparison with the other methods described in the literature. Finally, all the hyperspectral remote sensing registration methods developed in this thesis are compared.

Chapter 3 presents a Fourier-based method for the registration of two hyperspectral images called Hyperspectral Fourier–Mellin (HYFM). The method exploits the information contained

in the different bands of the images and is based on Principal Component Analysis (PCA), fractional Fourier transform (FRFT), multi-layer fractional Fourier transform (MLFFT) and a combination of log-polar maps. Two band reduction methods are evaluated. HYFM is compared with the other methods described in the literature in terms of the registration accuracy.

Chapter 4 details the implementation of HYFM on Graphics Processing Unit (GPU). Pseudocodes and optimisation strategies to exploit the GPU architecture are explained. The results are evaluated in terms of execution times and occupancy.

Chapter 5 presents a feature-based method, called Hyperspectral KAZE (HSI–KAZE), to register hyperspectral remote sensing images taking into account the spectral information. It consists of a trade-off solution between KAZE and Accelerated-KAZE (A-KAZE) for hyperspectral images, and it is based on a non-linear pyramidal scale space and Modified-SURF (M-SURF) descriptors. This chapter proposes a band selection method and a method for estimating the geometric transformation from the matched features; both were specifically designed for hyperspectral image registration. The results obtained under different images and different conditions are also discussed and compared with those of the other methods.

Chapter 6 describes the techniques and optimisation strategies to fully implement HSI–KAZE on GPU. The method is analysed by comparing the execution times to those of a parallel multi-threaded CPU implementation using OpenMP.

Chapter 7 presents a hyperspectral remote sensing registration method based on Speeded Up Robust Features (SURF). Hyperspectral SURF (HSI–SURF) is specially adapted to hyperspectral images by incorporating spectral information in different stages of the method. The algorithm is evaluated using different band selection methods as the first stage.

Chapter 8 presents Hyperspectral MSER (HSI–MSER), a feature-based method for registering pairs of hyperspectral remote sensing images. In contrast to the other feature-based methods developed in this thesis, HSI–MSER searches for common regions in both images. It is based on Maximally Stable Extremal Regions (MSER) as the region detector followed by Scale-Invariant Feature Transform (SIFT) as the feature descriptor. The method exploits the spectral information available in the images achieving excellent results in terms of registration accuracy and execution times.

In Chapter 9, a GPU scheme for the two-level registration of multispectral images is proposed. It is designed for efficient implementation in a multi-GPU system in which different scenes are registered in parallel on different GPUs. The first level consists of registering the different bands of each multispectral frame available for a scene. Registering the resulting frames is the objective of the second level. The scheme compares the results obtained by HYFM and HSI–KAZE.

Finally, Chapter 10 remarks on the contributions, presents the conclusions, and proposes some possible future research lines.

1.1 Motivation

Remote sensing is the acquisition of information about an object without having contact with it; i.e., the data are collected at a considerable distance. However, the term remote sensing is nowadays most used in connection with electromagnetic techniques of data acquisition about the Earth and other planets [26], and the methods used for their processing and interpretation. This is the meaning of the word used in this doctoral thesis.

A sensor mounted on a satellite, aeroplane, drone, or other types of aircraft, constructs an image capturing the energy reflected by the Earth. The term "energy" is most commonly used in this case as the electromagnetic spectrum from the low-frequency radio waves through the microwave, far infrared, near-infrared, visible, ultraviolet, x-ray, and gamma-ray regions. Depending on the source of the energy, we can talk of *passive* or *active* remote sensing. The *active* imaging systems capture the energy reflected by natural sources, for example, the sunlight or the proper energy that the Earth has itself (its own temperature) [123]. In the case of *passive* systems, the energy source is provided by the system itself, for example, by using laser [23] or radar [130]. These systems measure the energy scattered back to them.

The type of remote sensing data acquired depends on the type of information needed and, consequently, on the sensor used. In this doctoral thesis, we focus on images with high spectral and spatial resolution which are commonly referred to as multispectral and hyperspectral images. The most significant characteristic of these images is the wavelength range [13]. These images are collected by spectrometers that typically span the visible, near-infrared, and mid-infrared parts of the spectrum, i.e. from 0.4 of $2.5\,\mu\mathrm{m}$ [34]. They

consist of several contiguous spectral channels, also called bands. An image with 10 or fewer bands is called a multispectral image, and consequently, a hyperspectral image is an image with more than 10 bands, typically, more than 100. This implies that each pixel is a vector in which each component is a measurement of a specific wavelength [49]. Owing to this high spectral resolution, we can more accurately distinguish objects, plant species, and land cover classes, among others. The use of hyperspectral remote sensing images has been extended to a multitude of applications such as agriculture [44, 66, 70, 90], hydrological science [41, 46, 118, 125], geology [47, 48, 133, 144], quality control [22, 35, 54, 128], and land use classification [1, 11, 31, 131]. Hyperspectral remote sensing images provide global data from large areas of the Earth, which would not be possible to obtain in the same way on the ground, although this does not eliminate the need for ground information in some applications.

The advances in sensor development in the last decades allow obtaining hyperspectral remote sensing images at a lower cost than before. Hundreds of hyperspectral remote sensing images are taken every day owing to the increased availability of satellite remote sensing systems [20, 25, 61]. This allows us to obtain hyperspectral images of the same region of the Earth taken from different viewpoints at different times. The series of remote-sensing images are used in applications where it is essential to compare, study or find differences between images. Automatic change detection [32, 59, 77, 79], environmental monitoring [40, 120, 138], image fusion [7, 14, 63, 87] and super-resolution image creation [4, 68, 69], are examples of applications in which registration is a previous fundamental task [149]. The images must be registered before any further processing. Misregistration can produce large errors in the post-processing stages [21]. Moreover, image registration is used in all remote sensing system to refine the initial geolocation accuracy [71].

Image registration consists of aligning images of the same area which have been taken at different times, from different viewpoints, and possibly under different lighting conditions. The registration problem is formulated in this thesis as the alignment or registration of one image, called the reference image, with respect to a second image, called the target image. It involves locating and matching the same regions in both the images. The images usually present changes in objects, in illumination, spatial, spectral resolution, etc, and may not cover exactly the same region. The difficulty lies in estimating the geometric transformation that best maps one image to the other.

Remote sensing images have some characteristics that make their registration difficult and challenging as compared to other types of images, e.g. medical images [71]. Algorithms have to deal with a variety of sensors and data acquisition conditions; atmospheric and meteorological conditions; seasons that make natural elements such as trees, lakes, or rivers unreliable references for registration; different lighting conditions due to multitemporal effects such as the changes in the angle of the Sun during the year; and the huge amount of data. All of these complicate the alignment of remote sensing images.

Image registration can be done by hand. It is usually performed by selecting control points in both the images (assuming a two-image registration problem). The human, usually aided by software, marks distinctive points (control points) present in both the images. A geometrical transformation is then computed from these control points. The first drawback of manual registration is that it is a time-consuming task that becomes prohibitive for large sets of images. The second drawback is the difficulty of finding distinct points in some satellite images because of their image resolution or because of their possible correspondence to very monotonous urban or rural areas. There are images in which locating a landmark is relatively easy, e.g. the Cathedral in a remote sensing image of Santiago de Compostela or the Eiffel Tower in a remote sensing image of Paris, but in most remote sensing images, we do not have this type of reference point. Furthermore, it is common to mark points inaccurately or to not have control points distributed all over the image because of this lack of landmarks. Both the situations can lead to large registration errors.

The main goal of image registration research is to develop fully automatic algorithms in order to be able to process large amounts of data and to improve accuracy and efficiency, i.e. to solve most of the problems of manual registration. Image registration methods can be classified according to their nature in two categories [149]: area-based methods and feature-based methods. The first group, area-based methods, is based on the calculation of the differences in image pixel values [71]. They do not find the correspondence between the images looking for common points or regions; they directly exploit the image intensities without any structural analysis. Area-based methods can be subdivided into three categories: mutual information (MI) methods [15, 65], correlation-based approaches [9, 62], and Fourier transform methods [16, 122]. MI-based methods require that the intensities between images be stable, which is not a characteristic of remote sensing images. For this reason, these methods are widely used in medical imaging. Correlation-based methods involve the optimisation

of a similarity measure translating regions of the reference image to the target image. The similarity measure is computed for window pairs of both the images, which makes correlation-based methods very computationally expensive. In contrast, registration methods based on the Fourier transform are an efficient alternative for computing the correlation between images. Fourier-based registration methods rely on the Fourier shift theorem, which states that a circular shift in the spatial domain is equivalent to a phase ratio in the frequency domain. They are also a good choice when images have been taken under different conditions [71, 149]. These characteristics make them an interesting option in the case of hyperspectral remote sensing images. In general, area-based methods are computationally more efficient but their accuracy can be affected by noise, occlusions, large temporal changes, and other sources that may influence pixel differences.

In contrast, feature-based methods seek to detect distinctive features in objects or interest points at a higher level. The features are points, contours, regions, or lines that must have certain characteristics: invariant to geometric transformations, good localisation accuracy, and insensitivity to image degradation effects. These methods rely on the extraction of the same features in the images to be registered. On the basis of the knowledge of a number of corresponding features, a geometric transformation that aligns one image with respect to the other can be calculated. This high-level representation makes feature-based methods more resilient to illumination changes, intensity changes introduced by noise, and multisensor registration [53]. Feature-based methods are the most used in remote sensing images because these images are rich in details such as lines, edges, corners, and regions, which are susceptible to being detected. Four main stages may be identified in a feature-based method: feature detection, feature description, feature matching and registration. In the first stage, an interest operator is applied to the images for finding distinctive features. These features are usually extracted by building a pyramidal scale-space to achieve scale invariance. Thus, the images are summarised into a set of features. For example, if we are talking about points, a set of their coordinates. Once the features have been detected, in the second stage, a feature descriptor operator extracts the spatial information of each detected feature to build a unique identification called a feature descriptor. The unique identification or description is calculated taking into account the information of a reduced neighbourhood surrounding the detected feature. It is usually based on intensity information, local histograms of gradients, or geometric relationships between edges. The descriptor method has to be invariant to geometric transformations, illumination changes, noise, among others, in order to obtain almost the same descriptor in

the target images if they present rotations, translations or different scale factor. The third stage is feature matching in which features are matched between images. It involves the application of a similarity or distance measure to verify whether each pair of features is really a match or an outlier. The measure must be robust to avoid false matches that introduce registration errors when the geometric transformation is calculated. In the last stage, the matches are used together to compute a geometric transformation. As the distance or similarity distance used in the keypoint matching stage is not infallible, it is common not to use all the matches to calculate the geometric transformation. Matches are usually selected using parameter estimation methods, e.g. the Random Sample Consensus (RANSAC) [39] method.

Initially, area-based methods, particularly those based on the frequency domain because of their efficient computation, were designed only to deal with small displacements [71]. Chen *et al.* [16] modified the method incorporating a log-polar grid, thus extending it to large rotations and scales. The method was called Fourier-Mellin invariant symmetric phase-only matched filtering (FMI-SPOMF) and is one of the most popular Fourier-based registration algorithms. Since then, different Fourier-Mellin-based techniques have been proposed. Keller *et al.* [64] introduced the pseudopolar Fourier transform (PPFT) to improve the approximation of the log-polar map for the estimation of large translations, rotations, and scalings. The pseudo-log-polar Fourier Transform (PLPFT), which can be calculated quickly by utilising the Fourier separability property and the FRFT, was proposed in [76]. The PPFT was generalised in [19], while an octa-log-polar version of the PLPFT was proposed in [139]. A simpler approach was followed in [116], in which an adaptable MLFFT for image registration with lower interpolation errors in both polar and log-polar grids is proposed.

Among the feature-based methods, the SIFT [80] is the most popular. It tries to detect points with distinctive features called keypoints. The SIFT method consists of four stages: scale-space extrema detection, keypoint localisation, orientation assignment, and keypoint description. In the first step, a Gaussian scale space is created by performing Gaussian convolutions and interpolations. The original images are smoothed at different levels using Gaussian filters and at different scales. The result is a multi-resolution pyramid of the images. Thus, the invariant characteristic to image scaling is achieved. Thereafter, a difference of Gaussian (DoG) is obtained to detect the keypoints. A point is considered a keypoint if it is the local minimum or maximum as compared to its 26 neighbours (8 neighbours at the same scale and 9 at the adjacent scales) of the DoG scale space. Once candidate keypoints are obtained, the second

step carries out a selection to discard low-contrast keypoints, and their location and scale are accurately determined. In the third step, one or more consistent orientations are assigned to each keypoint to achieve the rotation invariance. These are calculated from the orientation histogram formed from the gradient orientations of points within the region of the keypoint. The size of this region depends on the scale and the gradient magnitude. Finally, in the last step, a 128 value descriptor for each keypoint is generated from a set of weighted histograms of the gradient orientation computed on different regions of the normalised neighbourhood. Then, the keypoints of each image are matched according to the ratio of the Euclidean distance between descriptors.

Since the publication of SIFT, different feature-based methods have been published following the same scheme. The SURF [6] was proposed to be computed considerably faster than SIFT. This method exploits integral images to build the scale space. This is an approximation of the Gaussian scale space that reduces the computation time. Regarding the descriptor, it is formed only by 64 dimensions and the orientation is estimated using the Haar wavelet. KAZE [36] proposed the use of non-linear diffusion filtering to build the scale space by applying selective blurring. A different approach is used in the Features from Accelerated Segment Test (FAST) method [124]. It implements a very fast corner detector based on machine learning.

However, all of these algorithms were developed to work with greyscale or RGB images, and not to deal with the spectral information available in the multi and hyperspectral images. Regarding area-based methods, few authors have proposed approaches to deal with the large amount of data of hyperspectral remote sensing images. In [82], a genetic algorithm as a search technique to estimate the optimum parameters for the matching of two multispectral datasets was used. This approach can be considered a fusion of the different decisions resulting from the registration of different bands. In [18], the researchers proposed the use of the expectation maximisation algorithm to solve the joint image registration and fusion problem. The registration parameters are estimated by maximising a conditional expectation function. In [8], a method for the automatic registration of multi-sensor remote imagery using multiple scales and resolutions is outlined. It introduces a self-organising network to control the movements of neighbouring tie points. However, none of these methods is based on Fourier transforms, the most efficient area-based method.

The same is true for feature-based methods. Some authors have proposed modifications

to SIFT for registering multispectral images. A scale restriction criterion to remove incorrect matching points in multispectral remote image registration was proposed in [143]. In addition, in [74], an orientation restriction criterion was proposed. A similar approach was followed in [85], in which a descriptor vector with four orientation bins instead of eight was implemented. These restrictions were also applied to SURF in [132]. Moreover, an approach that uses the spectral information of neighbouring keypoints in addition to the descriptor to discard false matches was presented in [58]. However, these methods only use one band of each image to perform the registration; i.e., they do not use all the available spectral information. Regarding hyperspectral registration methods, in [88], an interest point detector for hyperspectral images that uses the PCA to reduce the dimensionality is introduced. For each principal component (PC), a scale space is built. [24] presents a method where the scale space is created using a non-linear diffusion equation, taking into account the spectral information. The keypoint localisation involves the comparison of each keypoint with its neighbourhood according to its spectral signature. However, the evaluation of these methods is very limited as the datasets consist in most of the cases of images with small differences in scale or angle.

Recently, image registration algorithms for remote sensing images based on deep learning have also been proposed following the trend of using neuronal networks for remote sensing processing. In [142], features extracted with SIFT are combined with features extracted by a convolutional neural network in order to register remote sensing images. In [136], the feature extraction and matching stages of feature-based methods are replaced by the hidden layers and the output layer of a deep neuronal network, respectively. In [72], the researchers proposed a dual densely connected convolutional network to obtain the matching point displacement. The method is based on the estimation of the corner displacements of the images according to a regression model. However, deep learning registration approaches are limited to small scale factors and rotations [67]. Another limitation of these approaches is the lack of large labelled datasets to perform the training stage.

Registration, as well as all the usual tasks performed over multi and hyperspectral images, for example, change detection, classification or object detection, is computationally demanding. The main reason is the large amount of spectral information available in each image consisting of hundreds of spectral bands in many cases. Most hyperspectral registrations methods in the literature opt to ignore spectral information, and only deal with some bands, i.e. handling them as RGB images, in order to reduce computational requirements. The spectral

information provides the potential for more accurate and detailed processing. Thus, accuracy is sacrificed for computational performance. In real-time applications, e.g. disaster and damage control, reconnaissance and surveillance, or detection in a search and rescue scenario, the execution time becomes crucial, particularly now that the number of images available in space agency databases is continuously increasing. At the same time, new applications of hyperspectral images to solve new problems are emerging because of the affordable price of the sensors that are mounted over a variety of platforms, for example, unmanned aerial vehicles (UAVs). This situation makes it necessary to produce fully automated processing pipelines, particularly registration algorithms, which are efficient from the computational point of view and that efficiently exploit the spectral information to improve the registration accuracy.

In this thesis, the problem of developing faster and more efficient hyperspectral image registration is addressed. A Fourier-based method and different feature-based methods are implemented to align hyperspectral remote sensing images. To handle extreme registration situations, the algorithms efficiently exploit the available spectral information in several ways. Furthermore, they are projected onto a many-core GPUs enabling real-time applications even when a large amount of data is processed.

1.2 Hypothesis and objectives

This thesis addresses two main questions. One, whether it is possible to design more accurate multispectral and hyperspectral remote sensing image registration methods by exploiting the spectral information available in the different bands that form these images, and two, whether it is possible to develop efficient methods that can be executed in real-time on low-cost computing infrastructures.

These questions are formulated as the hypothesis of this thesis:

H1. *Exploiting the spectral information available in the different bands allows the alignment of multispectral and hyperspectral images with large and unknown scale factors and rotations.*

H2. *Registration of multispectral and hyperspectral images may be carried out in real-time by developing efficient and parallel algorithms for commodity hardware.*

To test these hypotheses, the following objectives are stated. The main aim of this thesis was *to develop more accurate and efficient multispectral and hyperspectral remote sensing image registration methods by exploiting the available spatial and spectral information*. In addition, to being able to register images in real-time applications or to process batches of images, another goal of this thesis was *to develop efficient algorithms in parallel architectures such as GPUs and multi-core architectures*. More specifically, the objectives were as follows:

O1. *Analysis of registration algorithms for greyscale and RGB images*. The aim was to study the state-of-the-art registration algorithms in order to incorporate in a later stage, the spectral information available in multi and hyperspectral images. We focused on area-based and feature-based registration algorithms. In the first group, we devoted special attention to methods based on the Fourier transform, and in the second one, we studied SIFT, KAZE, A-KAZE, SURF, and MSER, among others.

O2. *Analysis of band selection methods*. The redundancy of information among the bands in hyperspectral images is a well-known problem. The goal was to review the state-of-the-art in this area with the objective of reducing the dimensionality of hyperspectral images and keeping only the relevant spectral information.

O3. *Implementation of hybrid (spectral/spatial) registration algorithms*. The objective was to implement registration algorithms that exploited the spatial and spectral information

available in the multi and hyperspectral images in order to improve the accuracy of the alignment. The methods had to be able to register images with large and unknown transformations, and temporal changes, which are changes produced over time. The following methods are proposed:

1. Area-based methods. The HYFM is a method for the registration of two hyperspectral images based on the computation of Fourier transforms. The fast Fourier transform (FFT) is used to compute the cross-correlation (phase correlation) between the two images. The method exploits the information contained in the different bands of the images and is based on PCA, the MLFFT, and a combination of log-polar maps.

2. Feature-based methods. These methods try to detect a large number of common structures in both the hyperspectral images (HSIs) that will then be used to calculate the transformation to align the images. Three methods are implemented: HSI–KAZE and HSI–SURF, which are based on keypoint detection, and HSI–MSER, which is based on region detection. All of them exploit the spectral information available in the images by detecting features in several pre-selected bands as well as by including spectral information in the descriptor. Moreover, a band selection method called Entropy-Based Band Selection (EBS) is specifically developed to deal with the registration problem. An exhaustive histogram-based search method is also designed to estimate the final geometric transformation by selecting feature matches.

These three objectives, O1, O2 and O3, were addressed in the following articles:

- Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Fourier-Mellin registration of two hyperspectral images. *International Journal of Remote Sensing*, 38(11):3253–3273, 2017
- Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Alignment of hyperspectral images using KAZE features. *Remote Sensing*, 10(5), 2018
- Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. SURF-based registration for hyperspectral images. In *International Geoscience and Remote Sensing Symposium*, pages 63–66. IEEE, 2019

- Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Exploring the MSER-based hyperspectral remote sensing image registration. In *Image and Signal Processing for Remote Sensing XXVI*. SPIE, 2020
- Álvaro Ordóñez, Álvaro Acción, Francisco Argüello, and Dora B. Heras. HSI–MSER: Hyperspectral Image Registration Algorithm based on MSER and SIFT. Under review in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

O4. *Efficient implementation of the registration algorithms and methods on many-core GPUs by using CUDA*. The aim was to implement the developed algorithms and methods in parallel architectures to enable real-time execution or to efficiently process large amounts of data. HYFM and HSI–KAZE were projected on commodity GPUs by applying different optimisation strategies. Multi-core OpenMP implementations were also carried out to be used as a benchmark for performance comparison.

This objective was addressed in the following articles:

- Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. GPU accelerated FFT-Based registration of hyperspectral scenes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4869–4878, 2017
- Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Transformada de Fourier aplicada al alineamiento de imágenes multidimensionales en GPU. In Rafael Asenjo, Ángeles Navarro, Arturo González-Escribano, Diego R. Llanos, Sergio Cuenca Asensi, and Jesús González Peñalver, editors, *Jornadas SARTECO*, 2017
- Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. Towards a multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. In J. Vigo Aguiar, editor, *Computational and Mathematical Methods in Science and Engineering*, 2018
- Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. A multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. *Journal of Supercomputing*, 75(3):1551–1564, 2019

- Álvaro Ordóñez, Francisco Argüello, Dora B. Heras, and Begüm Demir. GPU-Accelerated registration of hyperspectral images using KAZE features. *Journal of Supercomputing*, 76(12):9478–9492, 2020
- Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU. In *International Geoscience and Remote Sensing Symposium*, pages 1575–1578. IEEE, 2021

1.3 Methodological tools

One of the objectives of this thesis was to develop efficient software to register multi and hyperspectral images in real-time by exploiting parallel architectures such as many-core GPUs and multi-core Central Processing Units (CPUs). Therefore, C/C++ was chosen to code the registration algorithms. CUDA was selected to exploit the parallelism of the GPU architecture and OpenMP for the CPU architecture, as will be explained in Section 1.3.1. Section 1.3.2 presents the commodity hardware, in which the experiments were carried out, and their main characteristics.

The software used (compilers, libraries, tools, etc.) is described in Section 1.3.3. The experiments were executed under Linux, and compiled using the GNU Compiler Collection (GCC) for the CPU implementations and the NVIDIA CUDA Compiler (NVCC) for the CUDA implementations. Moreover, additional libraries were used for specific computations to obtain the best performance.

Different multi and hyperspectral remote sensing images were used to evaluate the registration algorithms. They were captured by different sensors on different dates and under different viewpoints and illumination. The complete datasets are described in Section 1.3.4.

Finally, Section 1.3.5 explains the different measures used to assess the effectiveness, efficiency, accuracy and robustness of the registration methods developed as part of this PhD dissertation.

1.3.1 Parallel programming fundamentals

In this section, we introduce the main concepts behind the parallel programming model of OpenMP for the distributed shared memory on multiprocessors, and CUDA, a parallel computing platform and programming model that uses a GPU for general purpose computing.

OpenMP

The Open Multi-Processing (OpenMP) is the standard Application Programming Interface (API) for multi-threaded parallel programming on shared memory multiprocessors [105]. It consists of a set of compiler directives, library routines, and environment variables to control the runtime environment. The directives allow extending the C, C++, and Fortran languages to the OpenMP paradigm. OpenMP implements a type of parallelism called Single Program, Multiple Data (SPMD) that consists in splitting up and running a task simultaneously with

different input data in order to increase the performance. Thus, computer programs can be parallelised and executed efficiently on general-purpose CPUs.

An OpenMP program begins with a single thread which is called initial thread. The initial thread is executed sequentially until it encounters a parallel region. When this occurs, new threads are created along with an execution context for each. The new threads plus the initial thread divide the iterations of the parallel section between them; i.e., each thread executes a subset of the total number of iterations. The simplest example to illustrate this is a for loop. The execution context of each thread consists of its own stack for a safe execution without interfering with the stack of the other threads, and private and shared memory spaces for all the other program variables. The developer is responsible for indicating which variables should be shared among all the threads and which variables should be copied to each thread (private copy). Thus, threads have read and write access to their own variables but not to the variables of the other threads. In contrast, shared variables are located in a single storage location in the memory. They allow communication between the different threads and make coordinated computations possible. However, it is possible that multiple threads try to modify the same shared variable at the same time, or that some threads try to read a variable, while the others attempt to read it. These conflicting accesses are addressed using synchronisation directives such as a barrier, which is a point where each thread waits for all the other threads to arrive, or a critical section, which is a section of code that can only be executed by one thread at a time. At the end of the execution of the parallel region, there is an implicit barrier and all threads wait for the remaining threads to finish their executions. Once all the threads have finished execution, the initial thread resumes the execution of the sequential code until a new parallel region is encountered. This execution model is called the fork-join model.

The key factors that affect the performance of a program parallelised using OpenMP are coverage, granularity, load balance, locality, and synchronisation [12]. The concept of coverage is related to Amdahl's law. It is important to parallelise a large portion of our code to guarantee higher performance. The second key factor, granularity, is related to the cost of invoking a parallel section: the creation of new threads, barriers, synchronizations, etc. If the program invokes a large number of small parallel loops, the performance will be affected. By default, OpenMP divides a parallel section according to the number of iterations and the number of threads. It is called a static scheme. When some threads have to do more work than the others, we have an issue of load balance, the third factor, because some threads will finish earlier and will have to wait for the others. In such cases, it may be better to use a

dynamic scheme that schedules works on a "first come, first served" basis. The drawback of using dynamic schemes is that they can create synchronization and locality problems. These are the latter two factors that can affect performance. In these cases, using a static scheme with a small number of iterations per thread may be a solution. There is a trade-off between load balance and locality that the developer must analyse on a case-by-case basis.

CUDA

The Compute Unified Device Architecture (CUDA) is a parallel computing platform and API model that uses a GPU for general-purpose computing. It was created by NVIDIA in November 2006. Since then, the architecture and programming model has continued to evolve. Similar to OpenMP, CUDA is designed to work with C, C++, and Fortran.

The GPU provides higher instruction throughput and memory bandwidth than the CPU within a similar price and power envelope [99]. The GPU is designed to run thousands of threads in parallel, while the CPU, a few tens of them. The reason behind this is that GPUs are designed for a goal different from that of CPUs. GPU threads are simpler in terms of the type of task they can execute. GPUs devote more transistors to data processing than data caching and flow control, as is the case in CPUs. This increases the number of available threads for highly parallel computations because the GPU can hide latencies with computation. This design makes GPUs ideal for situations where large blocks of data have to be processed, and therefore, ideal for multi and hyperspectral image processing.

The NVIDIA GPU architecture is built around an array of Streaming Multiprocessors (SMs), each one containing many CUDA cores, caches, registers, etc., as shown in Figure 1.1. CUDA allows running programs using parallel functions called kernels [127]. Each kernel executes a set of parallel threads. Thus, each thread runs an instance of the kernel following the Single Instruction, Multiple Thread (SIMT) programming model; i.e., the threads execute the same instructions on different data. The programmer has to organise these threads into a grid of blocks (1D, 2D or 3D). A block is a set of threads that work together; i.e., blocks are independent of each other. This is illustrated in Figure 1.2.

Each SM executes these threads in groups of 32 parallel threads called warps. In other words, a warp is the number of threads that can run concurrently on an SM. To maximise the hardware utilisation and avoid latencies, it is necessary to have warps ready to execute the next instruction. This is closely related to the configuration (grid size) and resource requirements of the kernel, particularly registers and shared memory.

Figure 1.1: GP104 streaming multiprocessor diagram (92).

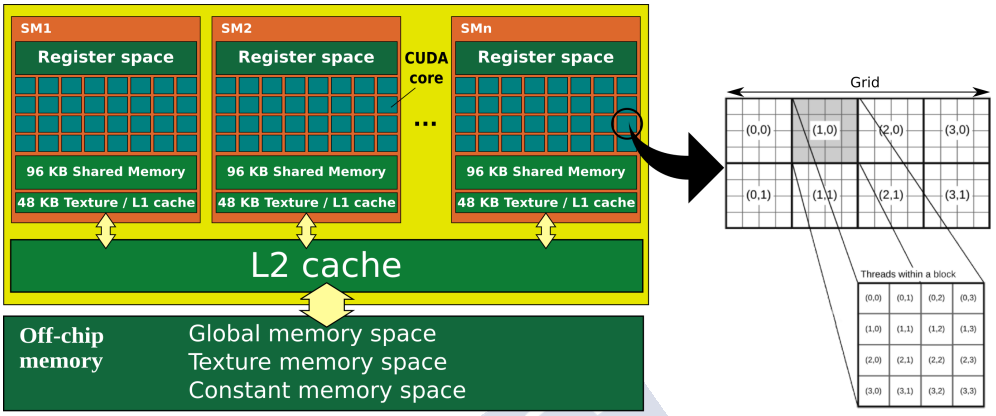


Figure 1.2: Overview of the block diagram of the GP104 GPU with a grid of blocks and a block of threads scheduled to one of the available SMs.

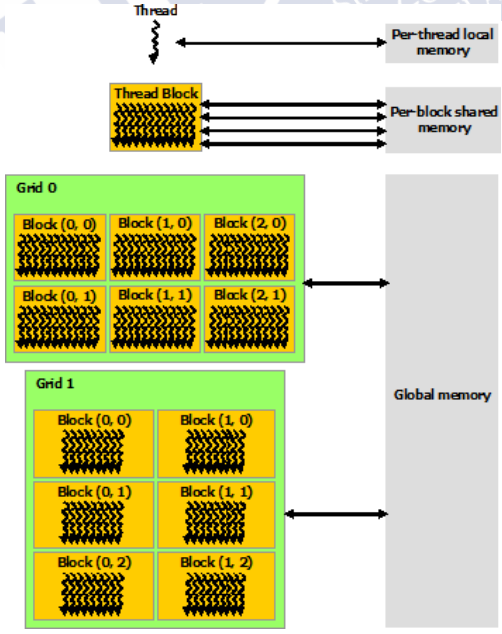


Figure 1.3: Memory spaces in CUDA (99).

Threads have access to different memory spaces, as can be seen in Figure 1.3. All threads have access to the global memory (DRAM). It is the largest memory space but has the lowest throughput. Other global addressable memories are the texture or constant memories which are read-only memory spaces for the thread but provide optimised accesses for some memory usages. The texture memory space is optimised for 2D spatial locality and improves performance, reducing the memory traffic. Therefore, it is efficient for interpolations. Finally, the constant memory space is used for storing data that do not change during execution. All of them are persistent across different kernel launches.

In contrast, the shared memory is only visible by the threads of the block and provides faster access to data than the global memory. Locally, threads have a set of registers and a private memory space. These spaces are not persistent across different kernel launches.

The developed methods were tested in the Kepler, Pascal, and Turing architectures, but we focused on the last two, the most recent ones during the development of this thesis. The Kepler architecture [94] includes a two-level cache hierarchy. There is 64 KB of on-chip memory for each SM, which can be configured as half each for the shared memory and the L1 cache, 48 KB of shared memory and 16 KB of L1 cache or vice versa. There is also a unified L2 cache of 1536 KB that is shared among all the SM units. In the Pascal architecture, the memory hierarchy is changed [95]. The Pascal architecture provides a fixed amount of dedicated shared memory and a unified L1/texture cache, which acts as a coalescing buffer for memory access. Thus, in Pascal, it is not necessary to select the preference for the shared memory and the L1 cache split to obtain the optimal performance in contrast to earlier architectures such as Kepler. For example, the GPU GP104 provides 96 KB/SM of dedicated shared memory and 48 KB/SM L1/texture cache (see Figure 1.2). Moreover, Pascal GP104 features a unified 2048-KB L2 cache. However, in the Turing architecture, NVIDIA again takes a similar approach to that of Kepler. The shared memory and the L1 cache are unified, and the programmer can again split the allocation of memory between both spaces depending on the workload.

The memory access pattern used has a considerable impact on the performance, particularly in the global memory. Therefore, it is important to design kernels that efficiently exploit the memory hierarchy, for instance, by performing fully coalesced memory access and minimising the data transfers between the global memory and the device by maximising the use of the on-chip memory (shared memory and caches).

In short, the developer has to take into consideration many aspects in order to yield the best performance. These considerations must be customised for each kernel.

Table 1.1: Main characteristics of the CPUs used in this thesis.

CPU model number	# of cores	# of threads	Base frequency (GHz)	RAM (GiB)	L1 (KiB)	L2 (KiB)	L3 (MiB)
Intel Core i5-6600	4	4	3.3	32	64 per core	256 per core	6
Intel Core i7-4790	4	8	3.6	24	64 per core	256 per core	8
Intel Xeon E5-2623v4	4	8	2.6	128	64 per core	256 per core	10

Table 1.2: Main characteristics of the GPUs used in this thesis.

	GeForce GTX 1070	Tesla P40	GeForce RTX 2070
Architecture	Pascal	Pascal	Turing
Compute Capability	6.1	6.1	7.5
GPU	GP104	GP102	TU106
Graphics Processing Clusters (GPCs)	3	6	3
Texture Processing Clusters (TPCs)	3	6	6
Streaming Multiprocessors (SMs)	15	30	36
Base clock (MHz)	1506	1303	1410
CUDA cores per SM	128	128	64
Tensor cores per SM	NA	NA	8
RT cores per SM	NA	NA	1
Device memory	8 GB	24 GB	8 GB
Memory clock (MHz)	2002	1808	1750
Shared memory per SM	96 KB	96 KB	96 KB
L1/texture cache per SM	48 KB	48 KB	
L2 cache	2048 KB	3072 KB	4096 KB
Texture Units per SM	8 units	8 units	4 units

1.3.2 Hardware

The methods developed in this thesis were implemented for their execution in different commodity hardware such as multi-core CPUs and many-core GPUs. The goal was was to ensure that a standard remote sensing user could use them using low-cost computing infrastructures and not need a supercomputer. We also tested the algorithms on low-end server hardware. The developed methods could be efficiently executed in the following hardware. The CPUs used in this thesis were Intel Core i5-6600, Intel Core i7-4790, and Intel Xeon E5-2623v4. The main characteristics of these CPUs are listed in Table 1.1.

With respect to the GPUs, the methods were evaluated on the GeForce GTX 1070 and GeForce RTX 2070, consumer-oriented GPUs for the PC gaming market, and the Tesla P40, a general-purpose GPU for high-end professional and scientific imaging applications. Table 1.2

Table 1.3: GPU resources defined by the compute capability.

Compute Capability	**6.1**	**7.5**
Threads/warp	32	32
Warps/multiprocessor	64	32
Threads/multiprocessor	2048	1024
Thread blocks/multiprocessor	32	32
Shared memory/multiprocessor (bytes)	98304	65536
Max. shared memory/block (bytes)	49152	65536
Register file size/multiprocessor (32-bit registers)	65536	65536
Max. registers/block	65536	65536
Max. registers/thread	255	255
Warp allocation granularity	4	4
Max. thread block size	1024	1024

shows their main characteristics: base block, number of SMs, CUDA cores per SM, device memory, cache sizes, etc. As can be seen in the table, the Tesla P40 has twice as many SM as the GeForce GTX 1070 at slightly lower clock speeds but with a larger L2 cache. However, the main feature of the high-end professional brand is the larger device memory, 24 GB in the case of Tesla P40 as compared to 8 GB for both the GeForce GPUs.

The GeForce GTX 1070 and Tesla P40 are based on the Pascal architecture and their compute capability is 6.1, while the GeForce RTX 2070 is based on the Turing architecture and its compute capability is 7.5. This number determines which hardware features and/or instructions are available on each GPU [99]. The general specifications of both the compute capabilities are outlined in Table 1.3.

As is to be expected, NVIDIA improves its GPUs with each new architecture. The GeForce RTX 2070 has twice as many SM as its predecessor, the GeForce GTX 1070, with a unified L1/texture cache and a double-sized L2 cache. Nevertheless, the big changes in the Turing architecture are not in the memory spaces. The Turing architecture features a new SM design, as can be seen in Figure 1.4. In Turing, two SMs are included per TPC, and each SM has a total of 64 FP32 cores, 64 INT32 cores, 8 Tensor cores, and 1 RT Core [93]. In comparison, the Pascal GP10x GPUs has one SM per TPC and 128 FP32 cores per SM. The new Tensor cores are specialised execution units designed specifically for performing the tensor/matrix operations typically used in deep learning. Tensor cores add new INT8 and INT4 precision modes and do not require FP16 precision. The new RT cores are dedicated to performing

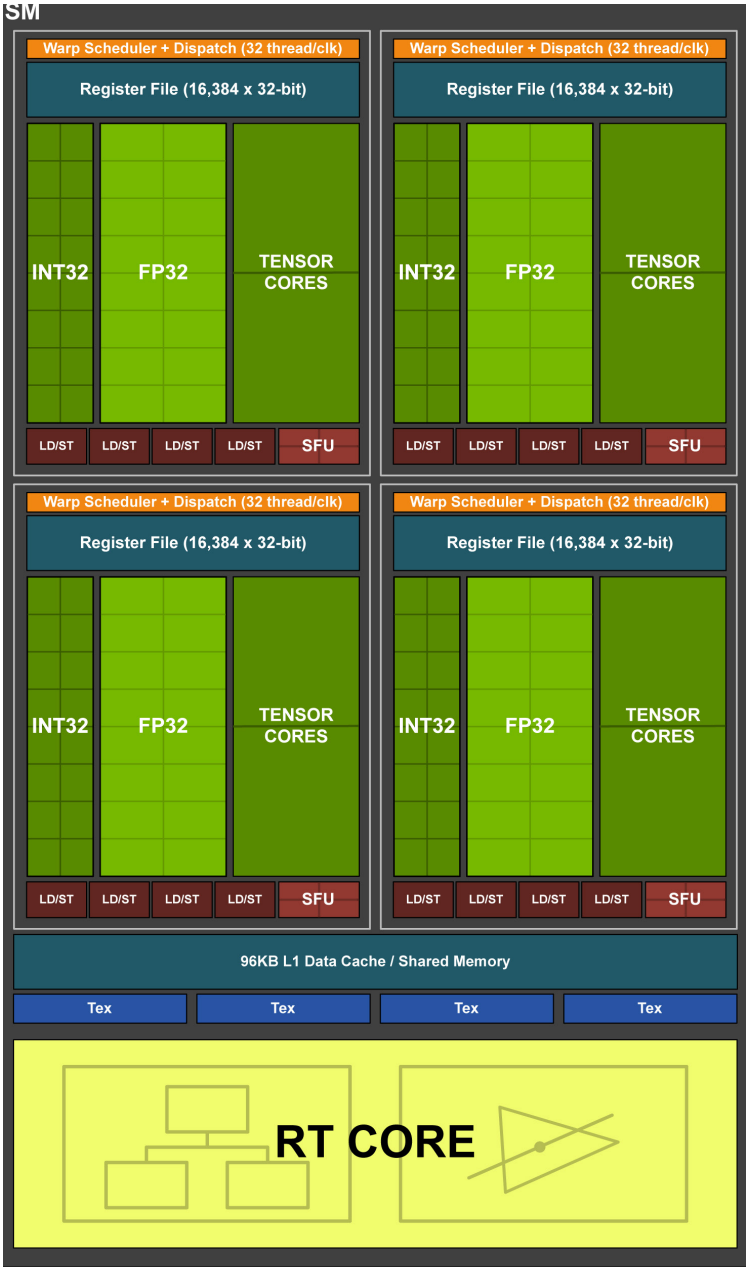


Figure 1.4: TU106 Streaming Multiprocessor Diagram (93).

ray tracing operations with extraordinary efficiency. The Turing architecture also adds INT32 cores that allow the GPU to execute FP32 and INT32 operations concurrently.

1.3.3 Software

All the methods developed in this thesis were coded using C and C++ programming languages. The codes were compiled using gcc and g++ [50] under Ubuntu with full optimisation flags -O3 in both the cases. Some complementary codes for processing the results and calculating the performance measures were coded using Python [43].

With respecto to the GPU implementations, the CUDA codes were compiled using nvcc [100] with full optimisation flags -O3. Diverse computing libraries have been implemented in the literature in order to obtain the best performance for image processing, linear algebra programs, and specific demanding computations such as FFT, among others. Most of them are included in the official NVIDIA CUDA Toolkit [100]. In this thesis, the following GPU-accelerated libraries were used:

- cuFFT [101]. This library was used to accelerate the computations of the FFT required in many stages of the area-based registration algorithm implemented in this thesis.
- Thrust [104]. It is a powerful library of parallel algorithms and data structures. In this work, it was used to perform GPU-accelerated sorting.
- cuBLAS [97]. It provides a GPU-accelerated implementation of the Basic Linear Algebra Subroutines (BLAS). cuBLAS was used in the PCA computation, which is done by the Eigenvalue Decomposition (EVD) of the covariance matrix of the images and in the feature-based registration algorithms to accelerate the feature matching stage.
- cuSOLVER [102]. This library provides a collection of dense and sparse direct linear solvers and Eigen solvers. It was also used in the PCA computation.
- NVIDIA Performance Primitives (NPP) [103]. It is a library of functions for performing GPU accelerated 2D image and signal processing. We used the NPP library to efficiently rotate and scale images.
- CUB [96]. CUB simplifies high-performance program and kernel development by taking care to implement the state-of-the-art in parallel algorithms. We used it to find the maximum and minimum values per band as well to compute histograms.

The NVIDIA Profiling Tool (NVPROF) was used to analyse GPU activity and optimise the code for best performance. Furthermore, OpenCV [42] and Python [43] were used to obtain the results of registration effectiveness for the state-of-the-art methods that we used to compare our methods.

1.3.4 Test images

The methods developed in this thesis were tested on remote sensing images taken by different sensors. The test procedure consists of registering one image, called the reference image, with respect to a second image, called the target image, which presents changes in scale, rotation, and translation. The test images could be classified into two datasets: hyperspectral and multispectral images.

Hyperspectral images

The first dataset consists of nine hyperspectral images taken by the Reflective Optics System Imaging Spectrometer (ROSIS) and Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensors. The ROSIS-03 sensor captures 115 bands in the spectral range from 0.43 to $0.86\,\mu\mathrm{m}$, corresponding mainly to the visible spectrum. The AVIRIS sensor captures 224 bands in the spectral range from 0.4 to $2.5\,\mu\mathrm{m}$, corresponding to the visible and infrared spectrum. Table 1.4 presents the detailed information for each image (the sensor, size, number of spectral bands, and spatial resolution). The complete dataset is available at `https://gitlab.citius.usc.es/hiperespectral/RegistrationRepository` [115].

This dataset can also be divided into two groups: images frequently used in the literature, for which only one image is available, and pairs of images taken by the AVIRIS sensor at different dates. The first group contains scenes of rural places and cities, and is commonly used for testing in the remote sensing community [55]. The set of target images is created by applying different scale factors and rotation angles to these images. The generation of the target images will be explained in more detail later.

Pavia University

It is an urban area surrounding the University of Pavia, Italy, taken by the ROSIS sensor and is made up of 103 spectral bands (the 12 noisiest bands were removed). The image size is 610×340 pixels with a spatial resolution of 1.3 m per pixel. Different structures are present in

Table 1.4: Sensor, size, number of spectral bands, resolution (m/pixel), and location of the test hyperspectral images.

Scene	Sensor	Size	Bands	Spatial Resolution
Pavia University	ROSIS-03	610×340	103	1.3
Pavia Centre	ROSIS-03	1096×715	102	1.3
Salinas Valley	AVIRIS	512×217	204	3.7
Indian Pines	AVIRIS	145×145	220	20
Jasper Ridge 2006	AVIRIS	1286×588	224	3.3
Jasper Ridge 2007	AVIRIS	1286×588	224	3.4
Santa Barbara Box 2013	AVIRIS	1024×769	224	15.2
Santa Barbara Box 2014	AVIRIS	1024×769	224	15.2
Santa Barbara Front 2009	AVIRIS	900×470	224	16.4
Santa Barbara Front 2010	AVIRIS	900×470	224	11.3
Baraboo Hill 2008	AVIRIS	1200×710	224	16.9
Baraboo Hill 2010	AVIRIS	1200×710	224	11.7
Crown Point 2010	AVIRIS	1400×540	224	3.5
Crown Point 2011	AVIRIS	1400×540	224	3.4

the image: buildings, asphalt, meadows, gravel, trees, and bare soil, among others. A colour composition of this image is presented in Figure 1.5(a).

Pavia Centre

It is a ROSIS image of the city of Pavia, Italy. Thirteen channels were removed due to noise, leading to an image of 102 bands with a spatial resolution of 1.3 m per pixel. It was originally 1096×1096 pixels in size, but a 381-pixel-wide black vertical band in the middle of the image was removed, resulting in a two-part image of 1096×715 pixels. Buildings, water, trees, meadows, soil, and asphalt, among other elements, are present in the image. A colour composition of this image is presented in Figure 1.5(b).

Salinas

It is a rural scene taken by the AVIRIS sensor in the Salinas Valley, California. This image has a size of 512×217 pixels with a spatial resolution of 3.7 m per pixel and 204 spectral bands (after the elimination of the water absorption and noisy bands). The scene includes

(a) Pavia University (b) Pavia Centre

Figure 1.5: Colour composition of the hyperspectral images Pavia University and Pavia Centre.

crops in different stages of growth (broccoli, lettuce, corn, and celery), stubble, bare soils, and vineyard fields. A colour composition of this image is presented in Figure 1.6(a).

Indian Pines

It was collected by the AVIRIS sensor over a rural area in NW Tippecanoe County, Indiana. It is a 220-band image (four noisy bands were removed) with a size of 145×145 and a spatial resolution of 20 m per pixel. The scene contains two-thirds agriculture (alfalfa, corn, grass, oats, soybean, and wheat) and one-third forest or other natural perennial vegetation. There are two major dual lane highways, smaller roads, a rail line, some low density housing and other built structures. A colour composition of this image is presented in Figure 1.6(b).

The second group consists of pairs of images of the same urban or rural scene taken at different dates by the AVIRIS sensor [61]. Thus, each pair of images presents changes in

(a) Salinas (b) Indian Pines

Figure 1.6: Colour composition of the hyperspectral images Salinas and Indian Pines.

Table 1.5: Reference geometric transformations for the second group of the test hyperspectral images.

Scene	Scale factor	Rotation angle	Translation (x,y)
Jasper	0.97	$-6.05º$	$(-9, 25)$
Santa Barbara Front	1.45	$3.52º$	$(2, 6)$
Santa Barbara Box	1.00	$0.00º$	$(32, 41)$
Baraboo Hills	1.44	$0.00º$	$(46, 132)$
Crown Point	1.03	$-0.70º$	$(-24, 48)$

infrastructure, vegetation, buildings, crops, and others, as well as, a different scale factor, rotation angle, and translation. The oldest image of each pair is used as the reference image and the most recent as the target image. The reference registration parameters are presented in Table 1.5, i.e. the expected outcome after registering these pairs of images. Moreover, as will be explained later, different scale factors, and rotation angles are applied to the target image of each pair in order to increase the experimental range.

(a) Reference image taken on 12 May 2006 (b) Target image taken on 13 August 2007

Figure 1.7: RGB images of the hyperspectral scene of Jasper Ridge.

Jasper Ridge

As their name implies, these images cover a region of the Jasper Ridge Biological Preserve, California. They include a forest, a lake, some buildings, vegetation, roads, and parts of bare soil, among others. The first image was taken in 2006 with a spatial resolution of 3.3 m per pixel and the second one in 2007 with a spatial resolution of 3.4 m per pixel. The second image has a drier appearance with respect to the first one. The size of the images is 1286×588 pixels and 224 spectral bands. The RGB images of this scene are presented in Figure 1.7.

Santa Barbara Box

These images were taken in a cultivation area near the city of Santa Barbara, California. The first image was taken in 2009 with clear weather and the second one in 2010 with a few clouds. Changes in crops are also presented. There have been changes in the crops from one

year to the next. Both the images have a size of 1024×769 pixels with a spatial resolution of 15.2 m per pixel and 224 spectral bands. The RGB images of this scene are presented in Figure 1.8.

(a) Reference image taken on 11 April 2013 (b) Target image taken on 16 April 2014

Figure 1.8: RGB images of the hyperspectral scene of Santa Barbara Box.

Santa Barbara Front

These images were collected over the coastal city of Santa Barbara, California. They include an urban area with numerous buildings, some vegetation, as well as the coastline. They are 224-band images with a size of 900×470. The first image was taken in 2013 and the second one in 2014. Both the images differ in spatial resolution, which is 16.4 m per pixel for the first image and 11.3 m per pixel for the second one. The RGB images of this scene are presented in Figure 1.9.

(a) Reference image taken on 30 March 2009 (b) Target image taken on 30 April 2010

Figure 1.9: RGB images of the hyperspectral scene of Santa Barbara Front.

Baraboo Hill

As the name indicates, these images were taken in Baraboo, Wisconsin. In the images, we can see a part of the city of Baraboo, as well as the Devil's Lake State. They also include forests, vegetation, crops, clouds, and roads, among others. The first image was taken in 2008 with a spatial resolution of 16.9 m per pixel and the second one in 2010 with a spatial resolution of 11.7 m per pixel. Both are 224-band images with a size of 1200×710. The RGB images of this scene are presented in Figure 1.10.

Crown Point

These images were taken in Crown Point, which is a rural neighbourhood located in Marrero, Louisiana. In addition to the urban area, they include vegetation, a river, canals, and roads. Both the images have a size of 1400×540 pixels and 224 spectral bands. The first

(a) Reference image taken on 13 July 2008

(b) Target image taken on 26 August 2010

Figure 1.10: RGB images of the hyperspectral scene of Santa Baraboo Hill.

image was taken in 2010 with a spatial resolution of 3.5 m per pixel and the second one in 2011 with a spatial resolution of 3.4 m per pixel. The RGB images of this scene are presented in Figure 1.11.

Multispectral images

The multispectral dataset consists of two scenes, called House and Reservoir [1]. Each scene consists of five and three different frames of 1280×960 pixels, respectively. The radiance bands, ranging from 475 to 842 nm., were captured by a MicaSense RedEdge MX sensor on 18 July 2018 in different UAV flights and, as a result, have different spatial resolutions.

Although the sensor captures all the bands at the same time, their co-registration is required. Figure 1.12 shows one frame of the House scene before co-registration. Fuzzy edges and the

[1]These images were obtained in partnership with the Babcock company, supported in part by the Civil Program UAVs Initiative, promoted by the Xunta de Galicia.

(a) Reference image taken on 30 September 2010

(b) Target image taken on 16 November 2011

Figure 1.11: RGB images of the hyperspectral scene of Santa Crown Point.

separation of the green, blue, and red channels are clearly observed, for example, in the roof and the small regular constructions. Once the bands are co-registered, the different frames of each dataset have to be registered. Figure 1.13 shows two co-registered frames from the House and Reservoir datasets before the alignment.

1.3.5 Performance measures

The methods developed in this thesis were evaluated in two ways: in terms of registration accuracy and in terms of computational times.

Performance results in terms of registration accuracy

The procedure followed for the experiments consists of registering one image, called the reference image, with respect to a second image, called the target image, which presents

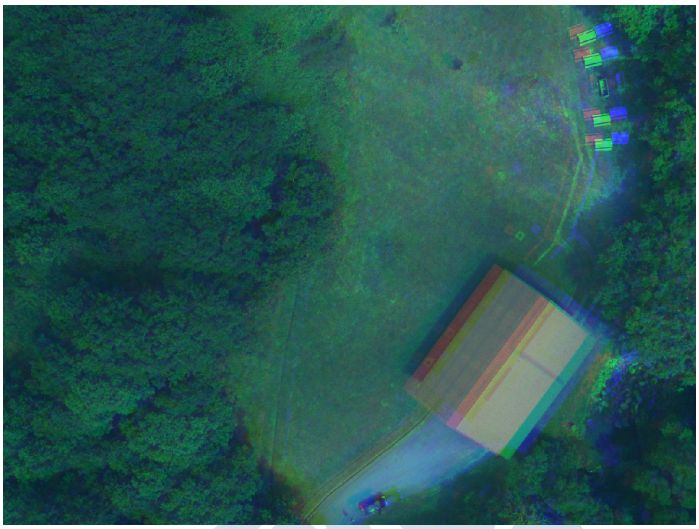
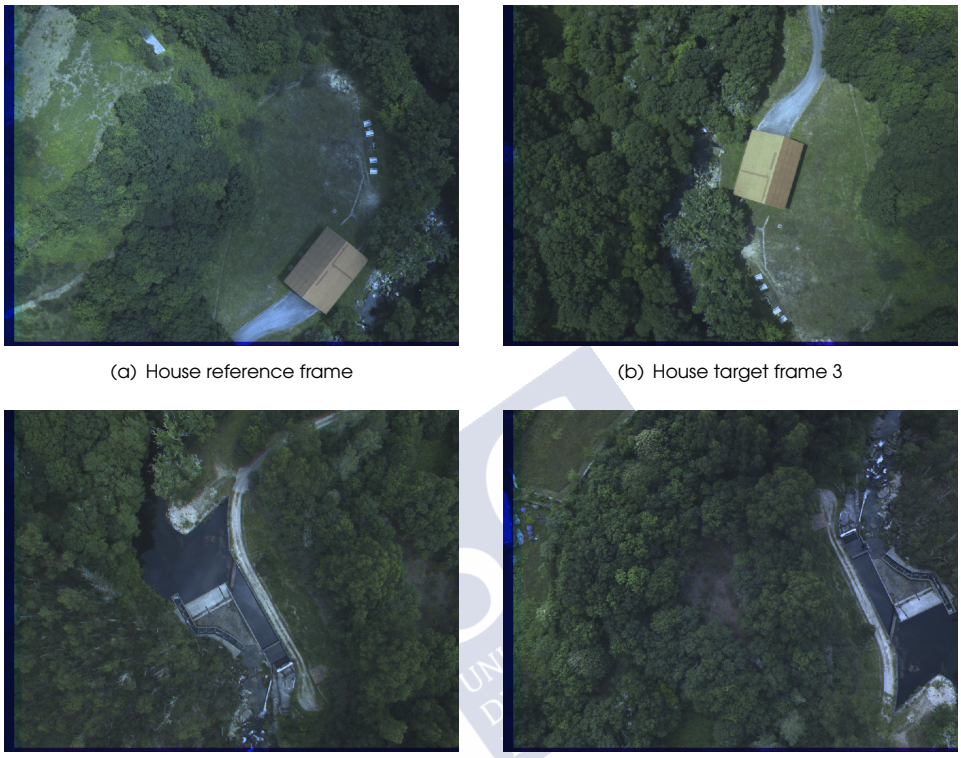


Figure 1.12: RGB composite of the reference frame of the House scene before co-registration.

changes in scale, rotation, and translation.

In order to evaluate the registration capabilities under extreme scaling and rotating conditions, and investigate all the registration details under the controlled conditions, the set of target images was extended. This extension was generated by applying a range of scale factors from $1/16\times$ to $25.5\times$ (65 scale factors) and rotation angles from 0º to 360º in increments of 5º (72 angles). From $1.0\times$ to $25.5\times$ the images are scaled up in steps of $0.5\times$, as shown in Figure 1.14, while from $1/2\times$ to $1/16\times$ the images are scaled down (scale factors of $1/i\times$, $i = 2, 3, \ldots, 16$). In the case of the hyperspectral images of the second group, these transformations were applied to the most recent image as mentioned in Section 1.3.4. In all the cases, the target images were trimmed in the central region to obtain the same size as that of the original images. The test consists of registering each target image (the generated ones) with respect to the reference image. As a result, $4,680$ cases were proved for each scene.

The range of scale factors that was correctly registered for the 72 angles for each scene was used as the main measure of registration accuracy. This could be summarised as the number of scale factors that were correctly registered. If an angle of a certain scale was incorrectly registered, i.e. a difference of 2.5º or more from the reference value, the entire scale factor was considered incorrect and the case was not included in the range. This measure was used



(a) House reference frame (b) House target frame 3

(c) Reservoir reference frame (d) Reservoir Target frame 2

Figure 1.13: RGB composite of some co-registered frames of the House and Reservoir scenes.

in all the articles that are part of this thesis.

Table 1.6 summarises the most common measures used to evaluate the registration methods in the literature. Therefore, the evaluations of the developed methods in this thesis are also presented according to them:

- **Chequerboard**. A chequerboard pattern was used to superimpose the reference and the registered target image in order to visually check the quality of the registration performed.

- **Number of matches**. This measure only makes sense when we are talking about feature-based registration algorithms. It is the number of matched features that then

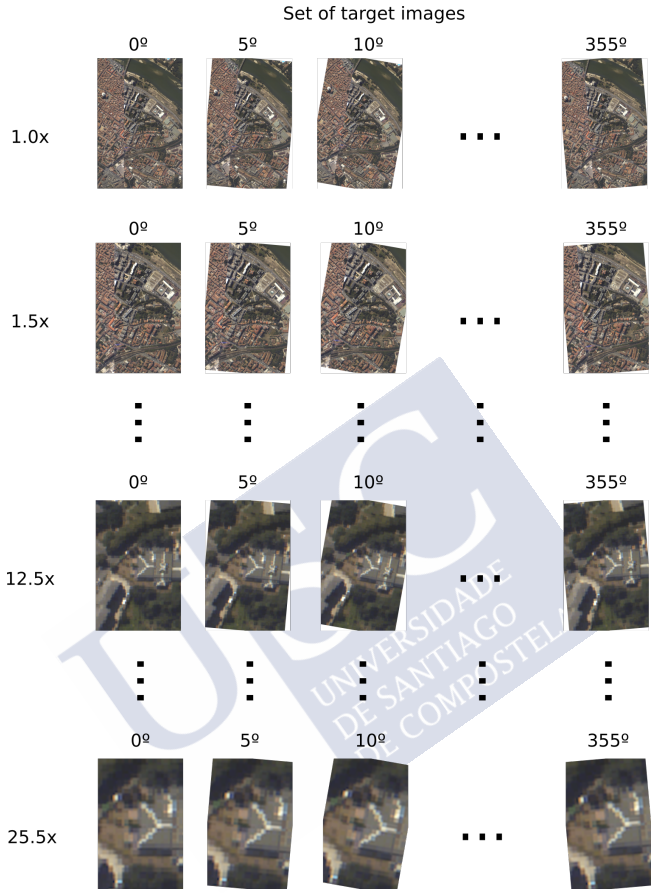


Figure 1.14: Method for creating a comprehensive set of target images by applying different scales and rotations.

are used to perform the registration. We need at least one match to register a pair of images. The number itself does not say much, we need to know how many of them can be classified as correct, i.e. how many have a low error according to the reference geometric transformations. What we know for sure is that a higher number of matches means longer computation time.

- **Registration error**. It could be calculated by using the extracted matches in the case of feature-based methods or manually selecting the control points (also called the reference

Table 1.6: Most common measures used to evaluate the effectiveness, accuracy, and robustness of the registration methods in the literature.

Ref.	Measures
[17]	Registration error, chequerboard
[18]	Registration error
[24]	Number of keypoints, correct match rate
[51]	RMSE
[52]	Chequerboard, correct match rate, number of correct matches, RMSE
[58]	Number of correct matches
[73]	Correct match rate, number of correct matches, RMSE
[74]	Chequerboard, correct match rate, RMSE
[81]	Chequerboard, number of correct matches, RMSE
[85]	Correct match rate, number of matches
[132]	Correct match rate, number of correct matches, number of matches
[140]	Number of matches, RMSE
[143]	Chequerboard, correct match rate, registration error
[146]	Chequerboard, number of correct matches, RMSE
[147]	RMSE

points) in the reference and the target images in the case of area-based methods. In both the cases, the registration error was computed as the average Euclidean distance between the features used in the registration of the reference image and the features used in the registration of the target image after the application of the reference transformation (see Table 1.5). It was measured in pixels.

Let $\vec{r}_i$ represent the feature i of the reference image, $\vec{t}_i$ its matched feature in the target image after the application of the reference transformation; and N, the number of matches,

$$\text{registration error} = \frac{\sum_{i=1}^{N} ||\vec{r}_i - \vec{t}_i||_2}{N}. \tag{1.1}$$

The features are the coordinates of the extracted keypoints or control points or, in the case of regions, the coordinates of the centre of the regions.

- **Root-Mean-Square Error (RMSE).** The RMSE was calculated as the square root of the second sample moment of the differences between the features used in the registration of the reference image and the features used in the registration of the target image after

the application of the reference transformation.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(||\vec{r}_i - \vec{t}_i||_2)^2}{N}} \tag{1.2}$$

- **Number of correct matches**. This is the number of correct matches among the total number. A match is considered correct if the registration error measured as the Euclidean distance between the feature of the reference image and the feature of the target image after the application of the reference transformation is lower than 2 pixels.

- **Correct match rate**. The quotient of the number of correct matches by the number of matches.

$$\text{Correct match rate} = \frac{\text{Number of correct matches}}{\text{Number of matches}} \tag{1.3}$$

Performance results in terms of computation time

The methods proposed in this thesis were also evaluated in terms of the computation time. They were implemented for efficient execution on many-core GPUs by using CUDA. In order to compare the GPU implementations according to a high performance baseline, multi-threaded implementations on a CPU were developed using OpenMP. Both versions are compared in terms of the computation time and the speedup. The procedure followed for the experiments consists of registering two images. The wall-clock time was provided for each test image and also is evaluated for each stage of the algorithms. It was defined as the time that a clock on the wall would measure as the elapsed time from the start to the end of the computation. We provide the average of ten independent executions for each experiment with and without the input/output access time. The reason for showing it both ways is that in many cases the algorithms are part of a processing pipeline in which the data are already stored in memory to be used as input in the next stage. Speedup measures the improvement in the execution speed between the GPU and the CPU implementations.

The achieved and theoretical GPU occupancy were also analysed for each stage of the algorithms. Occupancy is the ratio of the number of active warps per multiprocessor to the maximum number of possible active warps [98]. We obtained these measures by using the NVIDIA Profiling Tool (NVPROF), which provides the metrics at kernel level. Therefore, the occupancy values at the stage level were calculated by weighting each kernel occupancy

by its execution time with respect to the total execution time of the stage. A low achieved occupancy for a stage indicates that the GPU is not fully utilised because not enough thread blocks to hide the operation latency were launched. It is important to note that the kernels of library functions cannot be configured by the user and their occupancy is not always the best, and that higher occupancy does not necessarily imply higher performance. There is a point above which additional occupancy does not improve performance [98, 134].

1.4 Publications

The contributions of this doctoral thesis are included in the following publications:

1.4.1 International Journals

– Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Fourier-Mellin registration of two hyperspectral images. *International Journal of Remote Sensing*, 38(11):3253–3273, 2017.

Impact Factor (JCR 2017): 1.782.
Category: Imaging science & Photographic Technology. Rank: 11/27 (Q2).
Impact Factor (SJR 2017): 0.796.
Category: Earth and Planetary Sciences (miscellaneous). Rank: 54/278 (Q1).

– Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. GPU accelerated FFT-Based registration of hyperspectral scenes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4869–4878, 2017.

Impact Factor (JCR 2017): 2.777.
Category: Imaging science & Photographic Technology. Rank: 8/27 (Q2).
Impact Factor (SJR 2017): 1.547.
Category: Computers in Earth Sciences. Rank: 6/72 (Q1).

– Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Alignment of hyperspectral images using KAZE features. *Remote Sensing*, 10(5), 2018.

Impact Factor (JCR 2018): 4.118.
Category: Remote Sensing. Rank: 7/30 (Q1).
Impact Factor (SJR 2018): 1.430.
Category: Earth and Planetary Sciences (miscellaneous). Rank: 26/299 (Q1).

– Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. A multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. *Journal of Supercomputing*, 75(3):1551–1564, 2019.

Impact Factor (JCR 2019): 2.469.
Category: Computer Science, Theory & Methods. Rank: 31/108 (Q2).

Impact Factor (SJR 2019): 0.432.
Category: Hardware and Architecture. Rank: 64/148 (Q2).

– Álvaro Ordóñez, Francisco Argüello, Dora B. Heras, and Begüm Demir. GPU-Accelerated registration of hyperspectral images using KAZE features. *Journal of Supercomputing*, 76(12):9478–9492, 2020.

Impact Factor (JCR 2020): 2.474.
Category: Computer Science, Theory & Methods. Rank: 33/110 (Q2).

Impact Factor (SJR 2020): 0.445.
Category: Hardware and Architecture. Rank: 55/143 (Q2).

– Álvaro Ordóñez, Álvaro Acción, Francisco Argüello, and Dora B. Heras. HSI–MSER: Hyperspectral Image Registration Algorithm based on MSER and SIFT. The article is under review in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

1.4.2 International Conferences

– Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. Towards a multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. In J. Vigo Aguiar, editor, *Computational and Mathematical Methods in Science and Engineering*, 2018.

– Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Exploring the registration of remote sensing images using HSI-KAZE in graphical units. In J. Vigo Aguiar, editor, *Computational and Mathematical Methods in Science and Engineering*, 2019.

– Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. SURF-based registration for hyperspectral images. In *International Geoscience and Remote Sensing Symposium*, pages 63–66. IEEE, 2019.

Conference Rating (GGS Rating 2018): GGS Class 3 and GGS Rating B-.

Conference Rating (CORE 2021): Rank C.

– Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Exploring the MSER-based hyperspectral remote sensing image registration. In *Image and Signal Processing for Remote Sensing XXVI*. SPIE, 2020.

– Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU. In *International Geoscience and Remote Sensing Symposium*, pages 1575–1578. IEEE, 2021.

Conference Rating (GGS Rating 2018): GGS Class 3 and GGS Rating B-.

Conference Rating (CORE 2021): Rank C.

1.4.3 National Conferences

– Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Transformada de Fourier aplicada al alineamiento de imágenes multidimensionales en GPU. In Rafael Asenjo, Ángeles Navarro, Arturo González-Escribano, Diego R. Llanos, Sergio Cuenca Asensi, and Jesús González Peñalver, editors, *Jornadas SARTECO*, 2017.

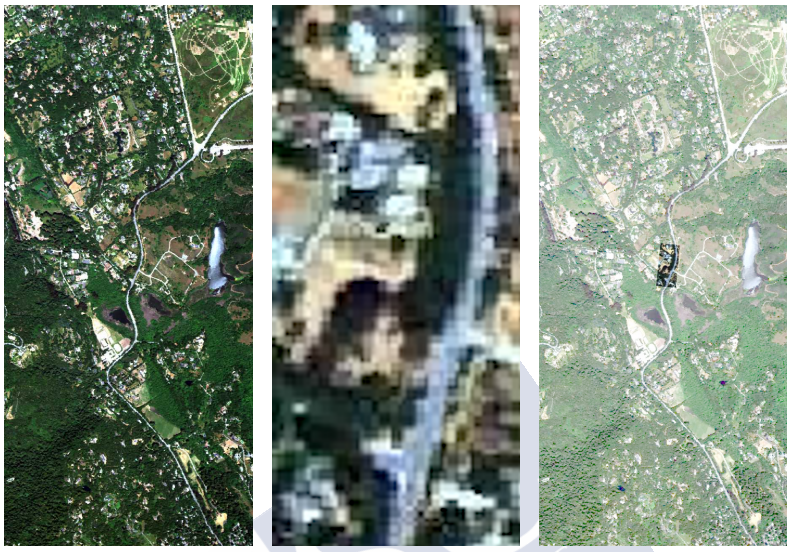
CHAPTER 2

DISCUSSION

Nowadays, hyperspectral images in remote sensing are used in a multitude of applications, helping to solve previously insoluble problems owing to the large amount of information that they provide [40, 77, 79, 120]. In applications such as land use activities, deforestation and climate change monitoring, disaster management, etc., image registration is a fundamental prior task in the processing pipelines.

As reviewed in Section 1.1, most registration methods are focused on RGB and greyscale images. Some methods have been proposed to register multispectral images [58, 74, 85, 132, 143], but they only use one band to register the images; i.e., they do not exploit all of the spectral information to improve the registration accuracy. The problem of multispectral registration is addressed in this way to avoid the high computational cost of processing all this information. The quick solution is to choose one band. In other articles [24, 88], methods for hyperspectral images using spectral information have been proposed. None of them deals with the problem of efficient registration for aligning large and/or constant batches of hyperspectral images or real-time applications.

The registration problem addressed in this thesis consists of finding the scale factor, the rotation angle, and the translation parameters, i.e. the similarity transformation which aligns two hyperspectral images fully exploiting the spectral information. An example illustrating the problem can be seen in Figure 2.1. As explained in Section 1.3.5, the methods developed in this thesis were developed to register scales factors of up to $25.5\times$ with arbitrary rotation angles. Table 2.1 summarises the scale factor range considered in the recent literature. In general, a small scale factor range is analysed in the literature, and in many cases, only one



(a) (b) (c)

Figure 2.1: Example of registration considered in this work: (a) reference image (size 1286×588), (b) target image (size 1286×588), and (c) result of the registration process showing the correctly registered superposition of the reference and the target registered image (scale $12.5\times$ and rotation angle $-5.77°$).

scale. The table includes methods for registering both panchromatic images and multispectral (from up to 31 bands) and hyperspectral (from up to 242 bands) images.

In light of the above, it is necessary to create efficient algorithms for hyperspectral remote sensing image registration. In this respect, different hyperspectral remote sensing registration methods capable of registering large scale factors are proposed in this thesis. We propose an area-based method called Hyperspectral Fourier–Mellin (HYFM) which is based on principal component analysis, MLFFT, a combination of log-polar maps, and peak processing. Moreover, an efficient registration processing pipeline was designed to support different feature-based methods. Three methods were developed taking this pipeline as the starting point: Hyperspectral KAZE (HSI–KAZE), Hyperspectral SURF (HSI–SURF), and Hyperspectral MSER (HSI–MSER). Furthermore, HYFM and HSI–KAZE, the most significant registration methods of this thesis, were projected onto a GPU for faster and more efficient hyperspectral image registration.

Table 2.1: Image type and the scale factor range considered by the registration methods in the literature.

Ref.	Image Type	Scale Factor Range
[17]	Multispectral	0.5× to 1.5×
[18]	Hyperspectral	1.0×
[24]	Hyperspectral	8.0×
[51]	Multispectral and hyperspectral	0.7×
[52]	Multispectral	1.0×
[58]	Multispectral	1.0×
[73]	Multispectral and hyperspectral	2.0×
[74]	Multispectral	0.5× to 2.0×
[81]	Multispectral	6.0×
[85]	Multispectral	1.0× to 2.0×
[132]	Multispectral	1.0× to 2.0×
[140]	Multispectral	1.0×
[143]	Multispectral	0.7× to 1.0×
[146]	Multispectral and hyperspectral	1.0× to 2.0×
[147]	Multispectral and panchromatic	2.8× to 4.0×

In this chapter, the different methods developed in this thesis are outlined and discussed. Finally, a final comparison is presented and analysed.

2.1 Area-based method

As stated in Section 1.1, registration area-based methods are those which deal directly with image intensities. They are commonly based on the correlation between images. An efficient way to compute this cross-correlation is to use the Fourier transform, particularly the fast Fourier transform (FFT) [16]. The Fourier-Mellin transform has been used as the basis for several registration methods. The best result reported using the MLFFT estimates scale factors up to $5\times$ in greyscale images [116]. Regarding the registration of hyperspectral images, the Fourier-Mellin transform has been used to register band-to-band hyperspectral images, i.e. not taking into account the spectral information [27, 28, 135]. Other methods have exploited this information [8, 18, 82], but they are intended for the recovery of small scale factors and rotations.

In this thesis, the first Fourier-Mellin-based method for the registration of large scale factors in pairs of hyperspectral images is proposed. Fourier-Mellin-based methods rely on

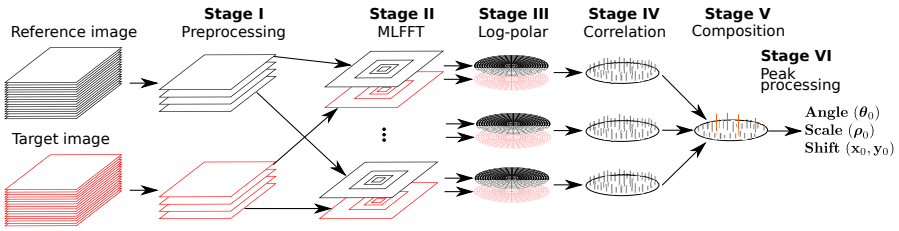


Figure 2.2: Schematic of the HYFM method for registration of two hyperspectral images.

the Fourier shift theorem, which states that a circular shift in the spatial domain is equivalent to a phase ratio in the frequency domain. The scale factor and the rotation angle are obtained by applying the phase correlation in a log-polar grid, while the translation parameters are recovered by computing the phase correlation in the Cartesian grid after correcting the scale factor and the rotation angle. It consists of six main stages, as can be seen in Figure 2.2. The stages are summarised below, but details can be found in Chapter 3:

- **Pre-processing stage**. In this first stage, a Blackman window and different filters are applied to remove frequencies that deteriorate the accuracy of the registration [57, 122]. In the processing of hyperspectral images, one issue is that they are highly correlated in the spatial and spectral dimensions. To reduce the dimensionality of each hyperspectral image and to retain as much as possible the variation present in them, a PCA computation is performed. Eight principal components (PCs) are retained. The registration is then applied to pairs of PCs.

- **MLFFT and log-polar stages**. The MLFFT technique is used for each PC to obtain a more robust log-polar map in the next stage [116]. It uses a multilevel grid to approximate the log-polar grid, particularly a four-level grid. Each level of the four-level MLFFT can be computed using the FRFT [5]. The FRFT is a generalization of the discrete Fourier transform (DFT) whose exponent is modified by a parameter α_L. The log-polar grid approximated by using these techniques is more accurate because the interpolation errors are reduced.

- **Correlation stage** In the fourth stage, the phase correlation is computed in the eight previous log-polar grids, one per pair of PCs. A matrix with approximate zeros everywhere except for some peaks is obtained. The scale factor and the rotation angle can

be obtained from the log-polar coordinates of these peaks. If the images to register are exactly equal and neither noise effects occur, only one peak should be obtained. However, this is not the case for pairs of remote sensing images taken at a different date. We have to analyse and choose one of these peaks (see the last stage).

- **Composition stage**. The performed experiments have proven that an efficient approach to exploit the spectral information, i.e. the eight pairs of PCs, is to combine the log-polar maps after the phase correlation. The resulting eight log-polar maps contribute independently to the cross-correlation by highlighting peaks with the correct registration parameters.

- **Peak processing stage**. As mentioned earlier, each peak of the resulting log-polar map after the phase correlation give us a possible scale factor and rotation angle. We have to choose the best one. To do so, the first 50 peaks are processed. This means that for each peak, a scale factor and a rotation angle are extracted. Both are applied to the target image; i.e., this image is rotated and scaled. The second phase of correlation in the Cartesian grid is computed using the reference image and the rotated and scaled target image. This process is repeated for each of the 50 peaks. The value of the highest peak in all the Cartesian grids will determine the correct peak in the log-polar map and the translation parameters for registering the images.

The results published in [106] (see Chapter 3) revealed that the method can register scale factors up to $7.5\times$ at arbitrary angles. The design of the method, as well as the various parameters mentioned, is the result of extensive experimentation. Different methods for dimensionality reduction have been proven for the pre-processing stage. PCA and Robust Colour Morphological Gradient (RCMG) were the two best performing methods. In the case of PCA, the use of different numbers of components were proven. In the case of RCMG, the dimensionality reduction is performed by calculating a gradient of the hyperspectral image on the basis of Euclidean distances among the pixel vectors by using the RCMG [29]. The result is an image with one band whose edges are enhanced and thinned. This one-band image is joined to the first four PCs. The idea is to supplement the PC information with an enriched band image with distinct information that helps to improve the registration.

As can be seen in the section results of Chapter 3, the configuration with eight PCs yielded better results than that of four PCs + RCMG. In the case of the images of *Jasper Ridge*, four PCs + RCMG failed to register them as no scale factor was correctly registered. The reason

was that the information retained by the RCMG is not sufficient for this complex urban scene. In contrast, the configuration with eight PCs correctly registered scale factors up to $3.0\times$ in this case and improved the results in other two scenes. Therefore, PCA was chosen as the method for dimensionality reduction in the preprocessing stage.

HYFM was also compared with other methods from the literature, both area-based and feature-based. From the first group, it was compared with the original Fourier-Mellin (FMI-SPOMF) [16] and with MLLFT [116]. From the second group, the best-known methods, for their good results, were selected, namely SURF [6] and two SIFT implementations: the original (SIFT-1) [80] and a more recent version (SIFT-2) [60]. As they were designed to align greyscale images and not deal with spectral information, they could not directly register hyperspectral images. Instead of choosing one band, which is a recursive solution in the literature, we used the first PC, which retained most of the variance over the hundreds of spectral bands. The proposed HYFM method provided the best results on average, correctly registering 11.43 cases versus 8.57 for SIFT-1, the second best method. Thus, the results proved our first hypothesis: the exploitation of spectral information improves the registration accuracy by increasing the number of corrected scale factors recovered. The good results obtained by SIFT-1 without any special adaptation to hyperspectral images led us to investigate feature-based methods.

In Section 2.3, we will discuss an efficient GPU implementation of the HYFM.

2.2 Efficient feature-based registration pipeline

Registration feature-based methods, and the best-known SIFT and SURF algorithms, were introduced in Section 1.1. They try to detect distinctive features such as points, regions or lines in both the images. On the basis of the knowledge of the location of the same features in both the images, a geometric transformation can be calculated to align them.

The big challenge in hyperspectral image registration is how to exploit the spectral information in an efficient way. As explained in Section 1.1, hyperspectral images provide us with valuable information to solve new global problems; they also present a large redundancy. Therefore, processing all bands is not a smart solution. The wise solution involves selecting those bands that have different information, in our case, different features.

In this thesis, we propose an efficient feature-based registration pipeline for hyperspectral images that can be used with any feature detector and feature descriptor with few adaptations. It exploits the spectral information available in the hyperspectral images by performing feature

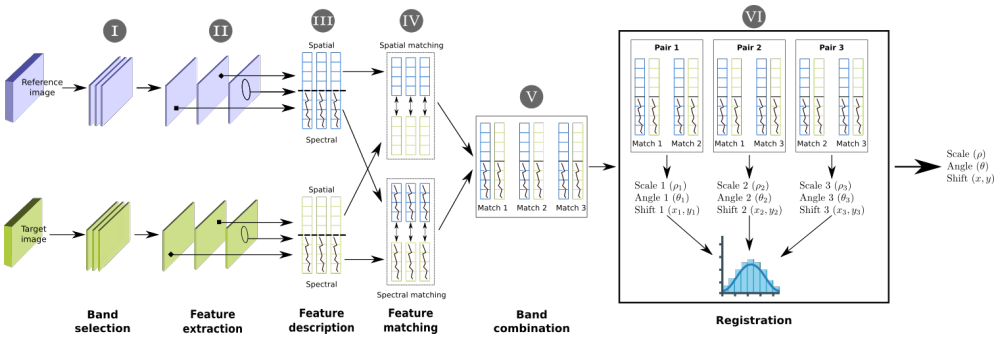


Figure 2.3: Schematic of the efficient feature-based registration pipeline for hyperspectral images. The feature extraction stage is illustrated with two types of features (points and regions).

detection and description in several bands of the images. Therefore, a band selection method focused on the hyperspectral image registration problem was designed and compared with the other methods described in the literature. Moreover, the registration pipeline incorporates spectral information into the feature descriptor. Thus, the false matched features are discarded in the matching stage. Furthermore, an exhaustive search method based on histograms is proposed to calculate the final geometric transformation by selecting some of the matched features, i.e. to perform the registration. Our approach achieved better results than the RANSAC method [39]. Figure 2.3 outlines the six stages that make up the pipeline. These stages are summarised below, but details can be found in Chapters 5, 7, and 8:

- **Band selection**. We propose a band selection method called EBS. The method involves selecting the bands with the highest entropy, but at the same time, they must be separated by the maximum number of consecutive bands. Eight bands are selected.

- **Feature extraction**. In this stage, a feature detector method is applied to each band previously selected. The aim of this stage is to detect a large number of common structures in both hyperspectral images that will then be used to calculate the transformation to align the images. In this thesis, we adapted the keypoint detection methods KAZE [36], A-KAZE [30], and SURF [6], as well as the region detector method MSER [84] to deal with remote sensing hyperspectral images.

- **Feature description**. In the third stage, the extracted features are described. The descriptor is made up of spatial and spectral information in order to achieve better differentiation. Implementations of the SIFT [80], SURF [6], or M-SURF [3] descriptors are used as the spatial descriptor. This spatial descriptor is enriched by a spectral part, particularly, the spectral signature of the feature.

- **Feature matching using spectral information**. Features of each pair of bands (one band for each hyperspectral image) are matched independently, i.e. without taking into account the features of the other pairs of selected bands. The matching is based on the distance between the spatial descriptors and the cosine similarity between the spectral signatures. The spectral information allows refining the matching process discarding outliers.

- **Band combination**. All the matches extracted from the selected bands are joined in this stage. Thus, features that are only present in some bands are used to compute the transformation; i.e., all the spectral information is considered together.

- **Registration**. Finally, we propose an exhaustive histogram-based search method to register the images. The method computes a possible transformation for each combination of two matches. A histogram is created using the angle values to make the first selection of possible transformations. The elements of the bin of highest frequency are sorted by scale factor to obtain the median. The scale factor, rotation angle, and translation parameters of the median element are selected to register the hyperspectral images.

Based on this pipeline, in this thesis, we propose HSI–KAZE (see Chapter 5), HSI–SURF (see Chapter 7), and HSI–MSER (see Chapter 8).

HSI–KAZE is a trade-off solution between KAZE and A-KAZE for HSIs and is based on a non-linear pyramidal scale space and M-SURF descriptors. The techniques used in SIFT and SURF to build the scale space do not respect the edges of the images, degrading important details at the same level as noise. KAZE proposes the use of a non-linear diffusion filter to build the scale space by applying selective blurring. In this way, KAZE avoids losing accuracy and distinctiveness in the keypoint detection stage. Furthermore, the noise is blurred, but the details and edges are maintained. Moreover, it uses the M-SURF descriptor, which is faster, and handles the boundaries better than the original SURF descriptor. In contrast to KAZE, A-KAZE uses a Fast Explicit Diffusion (FED) scheme to build the scale space. FED schemes are

characterised by their faster computation, better ease of implementation, and higher accuracy than those of the approach implemented by KAZE. Therefore, HSI–KAZE uses FED for the discretization of the diffusion equations.

SURF as the keypoint detector and descriptor was adapted to hyperspectral images according to this pipeline in [112] (see Chapter 7). As introduced in Section 1.1, SURF was created to improve performance with respect to SIFT. To do so, SURF opts to choose cheaper techniques in terms of computation times to build the scale space and the descriptor. In the construction of the scale space, SURF approximates the Gaussian convolutions with integral images and builds a pyramid of filters instead of a pyramid of images. Thus, SURF changes the scale of box filters rather than reducing the image size as in SIFT. With respect to the descriptor, it is based on the Haar wavelet and has only 64 components instead of Gaussian blurs and 128 components as in SIFT. This makes SURF a suitable solution to process large amounts of data, such as hyperspectral images, at a low computational cost. In this regard, we have proposed HSI–SURF.

A different approach is followed in HSI–MSER (see Chapter 8). It is a hyperspectral remote sensing registration method based on the detection of regions by using MSER and by describing them using an adaptation of the SIFT descriptor to the regions. MSER [84] extracts regions, called extremal regions, by thresholding the image at different grey levels and according to a stability criterion. In our proposal, regions are detected along with the selected bands and then transformed into ellipses. These regions are later described by a spatial and spectral descriptor. For the spatial part, SIFT is used. The SIFT descriptor is computed on the surface of the region bounded by the ellipse. HSI–MSER maintains for hyperspectral image registration the properties that make MSER an ideal method for greyscale image registration: it is unaffected by linear or affine transformations as well as geometric and photometric changes [86].

One of the main contributions of the pipeline is the proposed band selection method, i.e. Entropy-Based Band Selection (EBS). It has been compared with state-of-the-art feature reduction methods in Chapters 5, 7, and 8, particularly with the following methods:

- **Feature extraction methods**. They combine the different bands of the hyperspectral image into a small set of new features.
  - **Principal Component Analysis (PCA)** [123]. It is a well known statistical method to reduce the dimensionality in different types of problems. The main idea is to

eliminate the data redundancy present in the bands by means of a high correlation analysis. PCA generates a new set of linearly uncorrelated variables where the first few retain most of the variation present in all the original variables.

  - **BandClust** [10]. It is based on unsupervised clustering. It involves splitting the initial set of bands into disjoint clusters according to a mutual information criterion. Each iteration splits the original set of bands into two clusters. The method stops when the minimum for the criterion is found. In the end, the average of each cluster is computed to obtain the final reduced image.

- **Feature selection methods**. They do not modify the original data and are essentially limited to choosing a band or a set of distinctive bands.

  - **Ward's Linkage Strategy Using Mutual Information (WaluMI)** [83]. It performs an agglomerative clustering strategy. In the beginning, each band forms a single cluster. Next, the algorithm searches for the two clusters with the minimum dissimilarity difference. In order to do so, the dissimilarity matrix is built on the basis of the mutual information between each pair of bands. In the following steps, this matrix is updated using Ward's linkage strategy [137]. Finally, WaluMI chooses the most representative band of each cluster at the end. These bands are the final selection of the method.
  - **Entropy-Based Band Selection (EBS)**. Our proposal performs band selection on the basis of the entropy and the inter-band distance. In contrast to the other methods, EBS uses both hyperspectral images and is more computationally efficient.

Moreover, in the experiments described in Chapters 7 and 8, we added another case of study to the evaluation of feature reduction methods: performing the registration on a single band chosen at random. This is a recurrent solution in some registration methods in the literature. The test procedure is detailed in Section 1.3.5. The results for HSI–KAZE, HSI–SURF, and HSI–MSER demonstrated that EBS provided the best results. For example, in the case of HSI–KAZE, with the use of EBS 30.71 scale factors were correctly registered on average, while with WaluMI and BandClust only 28.43 and 25.71 scale factors were registered, respectively. In addition to the difference in the scales they managed to register, another drawback of these two methods based on clustering is the higher computational cost than that of EBS.

PCA provided the worse results among all the band selection methods considered, even worse than those of the random choice of a single band. The reason was that a different

transformation was applied to the reference and the target images, leading to a smaller number of mutual features present in both the images, or even none, recovered. Choosing a band at random was not the best option. Some structures are only present in some bands, and thus, we omitted using them to improve the registration problem. We once again validated our first hypothesis.

In the previous experiments, all the methods selected or extracted eight bands. The number was set experimentally for the case of EBS, except for BandClust in which the algorithm chose the number of bands to select. In Chapter 8, we evaluated how the number of bands used affected the registration effectiveness. In this experiment, a different number of bands using EBS was used, from 2 to 16 bands. The best results were achieved in the case of the use of 16 selected bands, but the difference was smaller than that in the case of the use of 8 selected bands. Using a large number of bands leads to higher computational costs. Therefore, we chose the EBS of 8 as the default configuration for the proposed algorithms.

Another important contribution is the exhaustive histogram-based search method for choosing the matched features and calculating the final geometric transformation. In Chapter 5, our method is compared to RANSAC [39]. The RANSAC algorithm was proposed in 1981 with a general purpose and not specifically to find the features that were matched with the lowest error in an image registration problem. Even so, it is the most widely used method of image registration for this purpose. It begins by calculating a geometric transformation by selecting a predefined number of matches, e.g. three matches. This first estimation of the geometric transformation is the model. Then, the remaining matches are checked to determine wheter they fit this model. The matches that fit well, i.e. have an error smaller than a pre-defined error, are included in a consensus set. If the consensus set has more elements than a pre-specified parameter, we check how good our speculated model is by calculating its error. This procedure is repeated N times. The model with the lowest error is the result. Our proposal is totally different. A detailed explanation can be found in Chapter 5.

In our experiments, RANSAC and our proposal were used in KAZE and A-KAZE in order to compute the geometrical transformation for different hyperspectral images. The results showed that the average number of successfully registered scale factors doubled for all pairs of images when the proposed methods were used; in particular, 9.00 and 8.29 scale factors on average wew correctly registered versus 4.29 and 3.86 obtained using the RANSAC version, respectively.

Furthermore, our proposals were compared with other methods described in the literature

following the procedure detailed in Section 1.3.5. In particular, they were compared with the original Fourier-Mellin (FMI-SPOMF) [16], HYFM, the original KAZE [36], and the original A-KAZE [30]. Our proposals, HSI–KAZE, HSI–SURF, and HSI–MSER, provided the best results. They could register images with scale factors of up to $24.0\times$ as compared to the $9.0\times$ achieved by A-KAZE. The exploitation of the spectral information allowed our pipeline-based methods to register higher scale factors with arbitrary angles. On the one hand, the detection performed on eight previously selected bands brought to the registration process more features that were only present in some of the bands. On the other hand, the spectral information included in the descriptor, as well as in the matching stage, allowed us to discard false matches, as can be seen in the experiments described in Chapter 5. It was not surprising that the methods implementing the proposed pipeline outperformed the feature-based ones described in the literature, as these did not exploit spectral information in any way. What was surprising was that they outperformed HYFM by a large number of scales. This will be discussed in Section 2.4.

As explained in Section 1.3.5, chequerboard, number of matches, number of correct matches, correct match rate, RMSE, and registration error are the most common measures used to evaluate registration methods in the literature. Therefore, our feature-based methods were evaluated according to these measures. The results confirmed that a higher number of matches were obtained by exploiting the spectral information, e.g. $21,437$ versus 23 for HSI–KAZE and A-KAZE for the *Jasper Ridge* images, respectively (see Chapter 5). This figure, in the case of our proposals, was refined by removing false matches after spectral discarding and eliminating repeated matches. Finally, $20,535$ and only 18 matches were classified as correct for this example, respectively. Thanks to the exploitation of spectral information, more correct matches were obtained. This large number of correct matches made it possible to register images with larger scale factors and rotation angles.

Regarding the registration error and RMSE, the results obtained for our proposals and the methods described in the literature were very similar, as shown in Chapter 5. However, it is important to note that these measures were obtained considering the original registration parameters; i.e., no additional transformations were applied to the images (see Section 1.3.5). In this respect, our proposals allowed the registration of a larger number of scales than the methods described in the literature, but we could not compare our methods with them under these conditions because the literature methods could not register these large scale factors.

More details can be found in the corresponding chapters on the methods which implement

this pipeline, i.e. Chapters 5, 7, and 8. In Section 2.4, all the methods developed as part of this thesis are compared with each other.

2.3 Efficient GPU implementations

GPUs have been demonstrated to be appropriate in many different problems for computationally efficient hyperspectral processing: classification [45, 117, 141], spectral unmixing [119], change detection [78], and segmentation [121]. In the literature, some registration algorithms for two-dimensional images in GPU have been proposed [33, 56, 89, 126, 129, 145, 148]. However, none of them has been applied to hyperspectral images, which contain more than a hundred times the data of a typical image, nor are they prepared to register remote sensing images, which are obtained in uncontrolled environments, unlike medical images, for example. Furthermore, the execution time is fundamental in real-time applications, e.g. disaster and damage control, and detection in a search and rescue scenario, among others, as explained in Section 1.1. In addition, the increased availability of satellites or other aircraft capturing hyperspectral remote sensing images makes efficient registration methods necessary.

In this thesis, we propose the first two implementations of hyperspectral remote sensing registration algorithms on a GPU to the best of our knowledge. The efficient implementations of HYFM, an area-based method, and HSI–KAZE, a feature-based method, are detailed in Chapters 4 and 6, respectively.

As mentioned, the HYFM is already very efficient in single-thread implementation because the FFT can be used to compute the cross-correlation between the images. Even so, for large images such as *Jasper Ridge*, more than 8 min is needed, which is far from a real-time execution. However, this algorithm is very suitable for implementation on commodity GPUs because the operations are mostly point by point and can be computed by independent threads, i.e. without data dependencies. More complicated is the GPU implementation of HSI–KAZE, in which the feature detection and feature matching stages are the most computationally demanding because of the complexity of the operations and the large amount of data involved. The algorithm is already computationally demanding in the CPU version, but it is the best performing method in terms of registration accuracy. Both the methods are implemented on a GPU without any loss in the registration accuracy, although some optimisation strategies have been applied by changing in some cases the method of computing some parts of the algorithms. In order to obtain the best performance and reduce the computational time, in addition to adapting the

method to the SIMT model, a set of optimisation strategies have been applied in both the algorithms:

- **Memory optimisations**.
  - **Reduce the data transfers between the host and device memories.** Data transfers are costly in terms of performance [98]. Implement algorithms that run entirely on the GPU to reduce computation times and to use the memory more efficiently. Data should be kept on the GPU as long as possible. Even if some kernel has low performance, it might be a good option not to transfer the data to the host and then to repeat the reverse operation afterwards.
  - **Use of shared memory whenever possible.** As explained in Section 1.3.3, a shared memory has higher bandwidth and lower latency than the global memory because it is on the chip. A shared memory is a good solution to improve memory access when, for example, there are data that all the threads in a warp will need.
  - **Use of texture memory whenever possible.** The texture memory is a dedicated read-only cache that includes hardware filtering. It can also perform linear floating point interpolation as part of the reading process. Compared with a typical CPU caching scheme, the texture memory cache is optimised for accelerating access patterns, for 2D spatial locality (in the coordinate system of the texture).
  - **Efficient use of the configurable L1 cache and shared memory size.** Some GPU architecture allows the configuration of the amount of shared memory and L1 cache. The configurable L1 cache/shared memory size allows the maximisation of the available amount of each type of memory for each block depending on the kernel requirements. Thus, the efficiency of the code is improved by choosing the best memory configuration for each kernel.
  - **Minimise the use of memory.** In order to minimise the global memory space required, some computations are performed in-place; i.e., the outputs are stored over the space previously allocated for the inputs.
- **Instruction optimisations**.
  - **Avoid divergences.** The number of flow control instructions is reduced to minimise divergences in the code, which negatively affects performance. A trivial case is

when the condition depends only on the thread ID. In this case, it is aligned with the warps and can be avoided.

  - **Use of low-level optimised instructions.** The use of low-level optimised functions provides better performance and better accuracy, particularly for code repeatedly executed [98].

- **Instruction and memory optimisations**.

  - **Use of vectorial instructions.** Vector loads and stores help to increase bandwidth utilisation and decrease the number of instructions. However, it is not a global solution for every kernel because it increases the register pressure.
  - **Use warp-aggregated atomics.** The use of atomic operations prevents the race conditions among threads but is clearly a bottleneck. Warp aggregation is the process of combining atomic operations from multiple threads in a warp into a single atomic [2]. This approach uses shared memory atomics so that the counter is only shared between the threads in the same block. This increases the performance and reduces the number of collisions in the global memory.
  - **Reduction of the number of operations using warp-level primitives to avoid shared memory latency.** Since CUDA 9.0, the API implements new warp-level primitives that make the data exchange in warp-level parallel reductions more efficient [75]. The reason is that the data are shared between the threads within a warp by using registers instead of a shared memory. These are commonly used in conjunction with warp-aggregated atomics.

- **General optimisations**.

  - **Search for the best kernel configurations to get the highest possible occupancy.** To hide latencies and keep the hardware busy, the maximum block size for each kernel is selected with the requirement that the number of registers and the shared memory usage do not act as the limiters of the occupancy [98].
  - **Change the order of operations to efficiently exploit the SIMT programming model.** The implementation must fit the architecture paradigm in order to achieve the best performance. Thus, parts of the algorithms sometimes have to be modified in order to efficiently exploit the GPU programming model, but always taking care to not affect the results of the registration.

– **Efficient computation using libraries.** Diverse computing libraries have been implemented in the literature in order to get the best performance for different specific tasks, as explained in Section 1.3.3.

The implementation details, as well as the pseudocodes, can be seen in Chapters 4 and 6. In both the cases, full GPU implementations have been realised and large accelerations are achieved as compared to CPU implementations.

For HYFM, 75% of the computation time is used for processing the steps that involve FFT computations. These steps are the MLFFT, the phase correlation and the high-pass filter. In them, the highest accelerations are achieved, i.e. speedups of up to $336\times$ are obtained as compared to a single-thread CPU implementation for the registration of the *Jasper Ridge* images. The use of the texture memory space is crucial to perform the large number of memory accesses and interpolation calculations required in the computation of the different levels of MLFFT. The FFT computations are accelerated using the cuFFT library [101]. Furthermore, image scaling and rotation is the operation that is executed many more times than the others. It is also computationally expensive because it requires arithmetic computations and involves interpolation calculations. In this step, a speedup of up to $42\times$ is achieved.

When the image size is large, better parallelisation and speedups are achieved: speedups of $269.7\times$ and $247.5\times$ and execution times of 1.88 s and 2.05 s are obtained for the *Pavia Centre* and *Jasper Ridge* images, respectively.

For HSI–KAZE, the GPU implementation is compared to an efficient OpenMP algorithm executed in the CPU. The highest speedup, $31.90\times$, is achieved in the keypoint matching and band combination stages. The approximation of the Euclidean distance by using matrix computations allows us to find the two-nearest keypoints for all the reference keypoints in parallel and efficiently. Two reasons are the key to this high computational efficiency. First, the matrix multiplication computed as part of this approximation is the reason for the high degree of parallelism achieved. Second, the use of a modified version of the insertation sort algorithm reduces memory usage. Furthermore, a significant acceleration is achieved in the keypoint detection stage, particularly $15.68\times$. It is achieved owing to the use of the proposed optimisation strategies, highlighting the change in the order of the operations in the keypoint location substage in order to efficiently exploit the SIMT programming model of the GPU, and the use of warp-aggregated atomics and warp-level primitives in parallel reductions in the scale space construction.

As in HYFM, the speedup achieved is better when the image size is larger. For example, a speedup of 13.53× and an execution time of 12.60 s are obtained for the registration of the *Jasper Ridge* images.

The achieved and theoretical average occupancies are analysed for both the algorithms. In both the cases, the figures show normal levels of occupancy. Both HYFM and HSI–KAZE on GPU were developed taking into account the three possible limiting factors in the occupancy:

- **Number of active threads per block**. The maximum number of active threads per SM depends on the compute capability supported by the GPU, as explained in Section 1.3.1. A block size and a number of threads per block that keeps the entire GPU busy should be chosen.

- **Number of registers per thread**. The maximum number of registers per SM and the maximum number of concurrent warps per SM depend on the compute capability supported by the GPU, as explained in Section 1.3.1. If a kernel requires too many registers per thread, the maximum number of simultaneous threads (number of concurrent warps × 32 threads per warp) will not be reached; i.e., we cannot hide latencies.

- **Amount of shared memory per block**. The maximum quantity of shared memory per block depends on the compute capability supported by the GPU. The shared memory is also partitioned between active thread blocks. If the amount of shared memory required per thread exceeds a certain value, the number of concurrent warps will be reduced.

However, we prioritise low execution time over higher occupancy because higher occupancy does not imply necessarily higher performance. For example, a lower occupancy kernel will have more registers available per thread than a higher occupancy kernel, which may result in less register spilling to local memory covering latency with low occupancy [98].

Thus, the results have proven our second hypothesis: registration of multispectral and hyperspectral images can be carried out in real-time by developing efficient and parallel algorithms for commodity hardware.

In the next section, both the methods are compared under the same hardware and on a multi-GPU system.

2.4 Final discussion

In this section, all the methods developed in this thesis are compared in terms of registration accuracy and execution times. This final discussion summarises the performance of the developed methods and puts their differences and possible applications into context.

Table 2.2: Comparison of the different hyperspectral remote sensing image registration methods developed as part of this thesis regarding the successfully registered cases for each scene. The number in parentheses summarises the number of scales that were correctly registered for all the angles. If an angle is incorrectly registered, the whole scale factor is considered incorrect; i.e., this case is not included in the table. The registration is considered correct if the parameters obtained by the algorithm are the same as the original values.

	Area-based method	**Feature-based methods**		
Scene	**HYFM**	**HSI-KAZE**	**HSI-SURF**	**HSI-MSER**
Pavia University	1/4× to 5.5× (13)	1/11× to 13.0× (35)	1/7× to 10.5× (26)	1/7× to 12.0× (29)
Pavia Centre	1/5× to 7.5× (18)	1/16× to 24.0× (62)	1/9× to 11.5× (30)	1/8× to 15.0× (36)
Indian Pines	1/2× to 4.0× (8)	1/4× to 5.5× (13)	1/3× to 5.5× (12)	1/3× to 8.0× (17)
Salinas	1/2× to 4.5× (9)	1/7× to 6.0× (17)	1/6× to 7.0× (18)	1/5× to 7.0× (17)
Jasper Ridge	1/5× to 3.0× (9)	1/12× to 12.5× (35)	1/2× to 4.0× (8)	1/3× to 7.0× (15)
Santa Barbara Front	1/4× to 3.5× (9)	1/9× to 9.5× (26)	1/6× to 4.0× (12)	1/3× to 7.0× (15)
Santa Barbara Line	1/4× to 6.0× (14)	1/12× to 8.5× (27)	1/4× to 3.5× (9)	1/5× to 7.0× (17)
Baraboo Hills	1/2× to 2.5× (5)	1/8× to 8.5× (23)	1/2× to 3.0× (6)	1/2× to 4.0× (8)
Crown Point	1.0× (1)	1/13× to 12.5× (36)	1/5× to 4.5× (12)	1/7× to 6.0× (17)
Average	9.56	30.44	14.78	19.00

Table 2.2 outlines the range of scales that were correctly registered for all the angles for all the developed methods in this thesis and for the entire hyperspectral dataset. HSI–KAZE provided the best results on average, correctly registering 30.44. The area-based method, HYFM, obtained the worst results; in particular, 9.56 scales were correctly registered on average for all the angles as compared to the 30.44 scales registered by HSI–KAZE. The large differences lay in the pairs of images of the same urban or rural scene taken at different dates from different viewpoints (*Jasper Ridge, Santa Barbara Front, Santa Barbara Line, Baraboo Hills, and Crown Point*). HYFM was proven to not be not very appropriate when there were large scale factors, displacements or changes in the images, as was the case. This was because local regions were deformed, so individual pixels aligned very poorly, but the derived features were more invariant to the changes and distortions [71]. In these situations, feature-based methods were more robust, as can be seen in Table 2.2.

As for feature-based methods, HSI–KAZE doubled the results of HSI–SURF, both based

on keypoint detection. The best results of HSI–KAZE was because of the robust and invariant scale space and descriptor used. HSI–KAZE built the scale space, in which the keypoints were detected, using a non-linear diffusion filter instead of Gaussian filters, used, for example, in SIFT. Thus, important structures in the images were preserved owring to the selective blurring of the diffusion. In contrast, HSI–SURF used SURF as the keypoint detector and descriptor, in which the original Gaussian scale space of SIFT was approximated by building a pyramid of filters instead of a pyramid of images and using the Hessian matrix. SURF sacrificed the good results of the Gaussian scale space for better computational performance. Therefore, HSI–SURF performs worse than HSI–KAZE.

HSI–MSER is also a feature-based method, but it differs from the other two in that it detects regions instead of keypoints, as discussed in Section 2.2. HSI–MSER provided the second best results on average: 19.00 scales with arbitrary angles were correctly registered. The spatial part of the descriptor used in HSI–MSER was based on SIFT, which performed better in terms of distinctiveness and invariance than SURF.

Table 2.3: Comparisons of HSI–KAZE, HSI–SURF and HSI–MSER with respect to the matches obtained to register the images. The last common scale successfully registered for all the methods is considered. HYFM is not included as it does not look for matches.

		HSI-KAZE	HSI-SURF	HSI-MSER
Jasper	Number of matches	606	33	31
	RMSE	**33.94**	122.95	73.23
	Registration error (in pixels)	44.47	52.30	**15.68**
Santa Barbara Front	Number of matches	1,538	94	134
	RMSE	17.64	154.32	**3.36**
	Registration error (in pixels)	7.00	76.50	**1.82**
Santa Barbara Box	Number of matches	342	143	85
	RMSE	179.33	199.87	**64.18**
	Registration error (in pixels)	91.00	111.33	**12.63**
Baraboo Hills	Number of matches	754	85	44
	RMSE	69.86	177.15	**13.24**
	Registration error (in pixels)	10.87	70.54	**4.43**
Crown Point	Number of matches	643	95	38
	RMSE	105.68	161.66	**3.20**
	Registration error (in pixels)	31.00	98.91	**2.84**

The RMSE and the registration error are measures frequently used in the literature to evaluate registration algorithms. To measure them, we use control points as explained in

Section 1.3.5. In the case of feature-based methods, the extracted matches are used as control points after applying the reference geometric transformation and, in the case of area-based methods, the control points are manually selected. Table 2.3 shows the number of matches, and their RMSE and registration error for the feature-based methods. The evaluations were performed using a common scale that the three methods could correctly register, particularly the last common scale successfully registered for all the methods for each image. As the scale increased, the methods got fewer matches, until there were not enough matches or quality matches to carry out an accurate registration. As HSI–KAZE is the method that manages to register the highest scales, it is also the method that obtains the highest number of matches on the last common scale. It increases the computation time, as mentioned earlier.

Regarding the RMSE and the registration error, HSI–MSER detects more accurate features. The images used to evaluate these measurements are the AVIRIS images taken at different times from different viewpoints and under different conditions. This suggests that region-based methods are more accurate than keypoint-based methods in images with great variability. However, HSI–KAZE correctly registers 30.44 scale factors on average, while HSI–MSER only 19.00 scale factors, as shown in Table 2.2.

Nevertheless, the matches are selected to avoid outliers and inaccurate matches in the final stage of the feature-based registration pipeline in order to then calculate the geometric transformation. Table 2.4 presents the numbers of matches used to compute the transformation, and their RMSE and registration error. In this case, the results are not as clear as before. HSI–KAZE and HSI–SURF perform better on two images each, and HSI–MSER in one. On average, HSI–SURF provides the best RMSE and registration error despite having the worst RMSE and registration before the selection (see Table 2.3). This confirms the considerable importance of the final stage and the good result obtained by our exhaustive histogram-based registration method.

The highest error occurs in the registration of *Crown Point*. These images, despite not having a large scaling and rotation, do have a non-linear distortion; i.e., it is not possible to obtain perfect alignment bu using a similarity transformation. More degrees of freedom are needed in the model.

The HYFM and HSI–KAZE methods were adapted to run on a multi-GPU system, particularly on two Tesla P40 (see Chapter 9). This multi-GPU registration scheme is proposed to co-register in parallel the bands of a set of multispectral images, and then to register the different scenes that form these multispectral images; i.e., two levels of registrations are performed.

Table 2.4: Comparisons of HSI–KAZE, HSI–SURF, and HSI–MSER regarding the matches used to register the images. HYFM is not included as it does not look for matches.

		HSI-KAZE	HSI-SURF	HSI-MSER
Jasper	Number of matches	2	4	4
	RMSE	**0.52**	0.94	1.50
	Registration error (in pixels)	**0.48**	0.94	1.30
Santa Barbara Front	Number of matches	4	2	3
	RMSE	5.99	**0.65**	0.75
	Registration error (in pixels)	5.98	**0.65**	0.75
Santa Barbara Box	Number of matches	2	4	4
	RMSE	2.06	**0.65**	1.41
	Registration error (in pixels)	2.06	**0.57**	1.33
Baraboo Hills	Number of matches	4	4	2
	RMSE	**1.00**	1.41	1.96
	Registration error (in pixels)	**0.88**	1.41	1.96
Crown Point	Number of matches	2	2	2
	RMSE	2.68	2.29	**1.91**
	Registration error (in pixels)	2.67	2.29	**1.84**

Figure 2.4 outlines the proposed two-level registration scheme that is executed on each GPU. Each GPU registers in parallel a set of multispectral images belonging to the same scene. The data are distributed to the GPUs by using OpenMP. Thus, two levels of parallelism are carried out. Moreover, bands and multispectral images are aligned by the two registration methods (HYFM and HSI–KAZE). At the end of the process, two alignments are obtained, one per each registration method. Thus, both the methods can be compared as the representatives of area-based and feature-based registration methods. This also demonstrates the versatility of the proposed methods. The methods and their efficient implementations on GPU are suitable to be part of a multi-level registration or a larger processing pipeline in which other tasks, e.g. change detection, are then applied.

To compare HYFM and HSI–KAZE, five control points were manually selected in each band and image of the multispectral dataset. In the first level, the registration of the bands of each multispectral image is performed. As can be seen in Chapter 9, the results show that HSI–KAZE registers the bands with an average RMSE of 1.38 pixels, while in the case of HYFM, the error reaches four pixels. As for the second registration level, in which the multispectral images are aligned to form a scene, HYFM does not manage to correctly register all the images. As explained in Section 1.3.4, these multispectral images were taken on

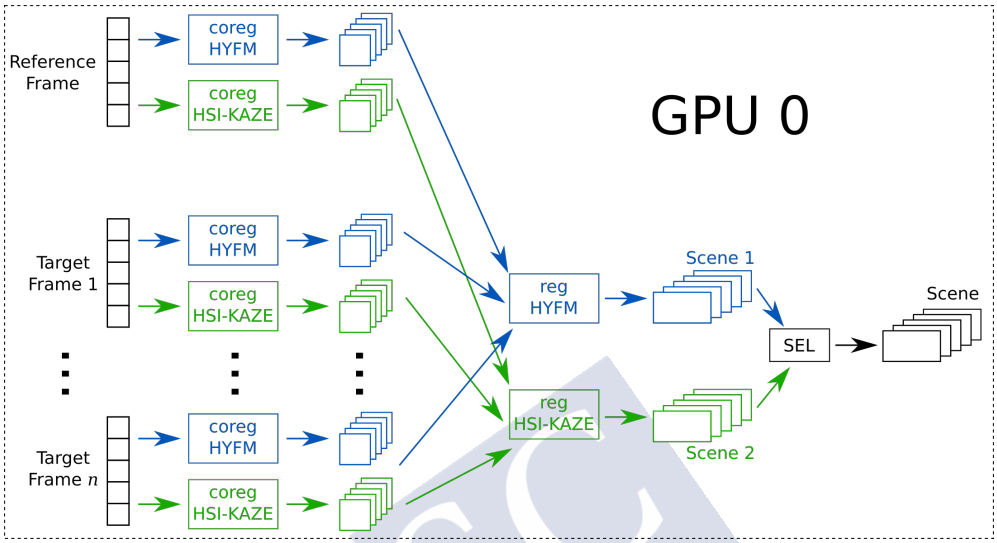


Figure 2.4: Schematic of the multi-GPU system for the registration of remote sensing multispectral images. Example for one of the GPUs (GPU 0).

different UAV flights and present changes in perspective, so that individual pixels align very poorly in the cross-correlation.

Regarding computational time, the procedure test consists of registering two hyperspectral images, as explained in Section 1.3.5. Table 2.5 shows the execution times in the CPU for each scene and method considering the last scale successfully registered in Table 2.2. This last point is not important in the case of HYFM. In contrast to the feature-based methods, HYFM always uses the same amount of time for each image; i.e., it does not depend on scale, rotation, or translation nor on the number of structures present in the images. HYFM is not based on detecting structures (features) such as HSI–KAZE, but on pixel correlation. The time required by HYFM depends solely on the spatial and spectral resolution of the images. Moreover, as explained before, Fourier-based methods are computationally efficient. All of these explain why HYFM achieves competent execution times as compared to feature-based methods.

Table 2.6 presents the execution times in the CPU per stage for the feature-based methods for the registration of the *Jasper Ridge* scene. In both the tables, we can see that HSI–KAZE is more costly than the other methods. There are two reasons.

The first is the number of features detected and then matched. The exhaustive search

Table 2.5: CPU execution times (in seconds) for each scene considering the last common scale successfully registered for all the methods. It includes CPU-GPU transfer times and the allocation of images in the memory.

	Area-based method	**Feature-based methods**		
Scene	**HYFM**	**HSI-KAZE**	**HSI-SURF**	**HSI-MSER**
Pavia University	29.08	91.56	7.37	15.99
Pavia Centre	127.96	748.37	36.55	28.97
Indian Pines	1.55	4.38	0.51	5.26
Salinas	6.47	21.15	3.41	6.58
Jasper Ridge	132.88	1,448.45	143.61	41.94
Santa Barbara Front	29.85	387.11	58.13	33.48
Santa Barbara Line	30.13	752.6	57.79	29.98
Baraboo Hills	135.54	1,730.19	252.93	38.69
Crown Point	146.29	2130.1	208.63	35.92
Average	71.08	812.66	85.44	26.31

to recover the registration parameters taking into account all the possible pairs of matches involves higher computational costs. Table 2.3 shows that the highest number of matches are detected by HSI–KAZE. The higher the number of matches is, the longer is the computation time at this stage. When a feature-based method fails, it is due to, among other things, a low number of matches or not enough accurate matches. This means that the number of matches is lower at higher scale factors. As explained before, HSI–KAZE is the method that manages to register the highest scale factor, and in this experiment, we considered the last common scale successfully registered, i.e. $3.0\times$ for *Jasper Ridge*, in order to compare all the methods under the same conditions. HSI–KAZE could register the *Jasper Ridge* images up to $12.5\times$ owing to the superior number of matches and the accuracy of these matches, while HSI–SURF and HSI–SURF only achieved registrations up to $4.0\times$ and $7.0\times$, respectively. As a result, it was expected that HSI–KAZE would have more matches at $3.0\times$ and present a longer execution time.

Second, the feature detection and description in HSI–KAZE are performed by constructing a non-linear pyramidal scale space. This approach allows more and most robust invariant matches to be obtained, as demonstrated by the results of Table 2.2, in which HSI–KAZE achieves registrations of up to $24\times$. However, it is computationally more demanding than the approaches used in HSI–SURF and HSI–MSER, as can be seen in Table 2.6. SURF was

specially created to be more computationally efficient at the cost of worsening quality. This makes HSI–SURF a suitable solution when the speed of computation is more important than the quality of the registration. In contrast, HSI–MSER implements a completely different feature detection stage. The method extracts a number of regions, called MSERs, by thresholding the image at different grey levels and according to a stability criterion. Then, an MSER can be defined as a stable connected component through different grey levels of an image. The connected components of the level sets can be organised in a tree structure. Thus, the complexity of computation of the detection stage is reduced to quasi-linear by using an approach based on the union-find procedure [91]. However, as the spatial descriptor in HSI–MSER, we used an approach based on the SIFT descriptor. The higher complexity of the SIFT descriptor explains the higher computational cost in this stage than that of HSI–SURF.

The execution times of the other two stages, band selection and registration, were almost the same because our efficient proposals for both the stages were used in all the methods. The higher number of matches extracted by HSI–KAZE resulted in a small increase in the execution times for the registration stage than can be seen in Table 2.6.

Table 2.6: CPU execution times (in seconds) per stage of the proposed efficient feature-based registration pipeline for the registration of the *Jasper Ridge* images considering the last common scale successfully registered for all the methods, i.e. $4.0\times$. It does not include the allocation of images in memory.

	HSI-KAZE	**HSI-SURF**	**HSI-MSER**
Band selection	0.40	0.40	0.40
Feature detection	484.38	11.46	10.52
Feature description	94.58	6.99	26.39
Feature matching	868.44	125.30	3.67
Band combination			
Registration	0.20	0.01	0.03
Average	1,448.00	144.16	41.01

As discussed in Section 2.3, HYFM and HSI–KAZE were implemented to perform efficient parallel registrations on GPU. In Chapter 9, both the methods are compared for the registration of bands and multispectral images, while in Chapters 3 and 6, both the GPU implementations are presented and compared with their corresponding CPU version. Table 2.7 presents the execution times for both the methods and for the registration of the whole hyperspectral dataset considering the last common scale successfully registered by both. In contrast to the articles

in which each implementation was presented, both the methods were run on the same GPU, on the GeForce RTX 2070, whose characteristics were described in Section 1.3.1. The execution times were reduced owing to the exhaustive parallel implementations that efficiently exploited the GPU architecture. On the GPU, HYFM required 1.38 s on average, while HSI–KAZE needed 4.10 s. Both the methods on GPU could be used to register hyperspectral images in real-time situations or to process long datasets in a reasonable amount of time.

Table 2.7: RTX 2070 GPU execution times (in seconds) for each scene considering the last common scale successfully registered for HYFM and HSI–KAZE. It includes CPU-GPU transfer times and the allocation of images in the memory.

Scene	**HYFM**	**HSI-KAZE**
Pavia University	0.49	0.71
Pavia Centre	1.66	3.48
Indian Pines	0.15	0.18
Salinas	0.32	0.39
Jasper Ridge	2.34	6.96
Santa Barbara Front	1.03	2.87
Santa Barbara Line	1.63	4.47
Baraboo Hills	2.50	9.48
Crown Point	2.34	8.32
Average	1.38	4.10

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 3

FOURIER-MELLIN REGISTRATION OF TWO HYPERSPECTRAL IMAGES

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.1080/01431161.2017.1292071, or with this information:

> Álvaro Ordóñez [1], Francisco Argüello [2], and Dora B. Heras [1]. Fourier-Mellin registration of two hyperspectral images. *International Journal of Remote Sensing*, 38(11):3253–3273, 2017

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 4

GPU accelerated FFT-based registration of hyperspectral scenes

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.1109/JSTARS.2017.2734052, or with this information:

> Álvaro Ordóñez [1], Francisco Argüello [2], and Dora B. Heras [1]. GPU accelerated FFT-Based registration of hyperspectral scenes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4869–4878, 2017

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 5

Alignment of hyperspectral images using KAZE features

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.3390/rs10050756, or with this information:

> Álvaro Ordóñez [1], Francisco Argüello [2], and Dora B. Heras [1]. Alignment of hyperspectral images using KAZE features. *Remote Sensing*, 10(5), 2018

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 6

GPU ACCELERATED REGISTRATION OF HYPERSPECTRAL IMAGES USING KAZE FEATURES

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.1007/s11227-020-03214-0, or with this information:

> Álvaro Ordóñez [1], Francisco Argüello [2], Dora B. Heras [1], and Begüm Demir [3]. GPU-Accelerated registration of hyperspectral images using KAZE features. *Journal of Supercomputing*, 76(12):9478–9492, 2020

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.
[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.
[3]Faculty of Electrical Engineering and Computer Science, TU Berlin, Germany.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 7

SURF-BASED REGISTRATION FOR HYPERSPECTRAL IMAGES

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.1109/IGARSS.2019.8900462, or with this information:

> Álvaro Ordóñez [1], Dora B. Heras [1], and Francisco Argüello [2]. SURF-based registration for hyperspectral images. In *International Geoscience and Remote Sensing Symposium*, pages 63–66. IEEE, 2019

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 8

HSI–MSER: Hyperspectral Image Registration Algorithm based on MSER and SIFT

The following is a reproduction of an article for which the author of this thesis is the main contributor. The article is under review in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. The preprint version can be found online under the following DOI 10.5281/zenodo.5494413, or with this information:

> Álvaro Ordóñez [1], Álvaro Acción [1], Francisco Argüello [2], and Dora B. Heras [1]. HSI–MSER: Hyperspectral Image Registration Algorithm based on MSER and SIFT.

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 9

Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU

The following is a reproduction of an article for which the author of this thesis is the main contributor. This is a verbatim reproduction, and the original can be found online under the following DOI 10.5281/zenodo.5494347, or with this information:

> Álvaro Ordóñez [1], Dora B. Heras [1], and Francisco Argüello [2]. Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU. In *International Geoscience and Remote Sensing Symposium*, pages 1575-1578. IEEE, 2021

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Spain.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 10

Conclusions

In this thesis, the problem of developing faster and more efficient automatic hyperspectral image registration was addressed. The focus was on designing and developing registration methods that provide good registration results in terms of accuracy and lead to efficient computation in commodity hardware. A Fourier-based method and different feature-based methods are implemented to align hyperspectral remote sensing images with large and unknown initial transformations. To handle these extreme situations, the algorithms efficiently exploit the available spectral information in different ways. Furthermore, they are projected onto many-core GPUs enabling real-time applications and the processing of large amounts of data.

During the development of this thesis, several problems or difficulties have been encountered. There is neither a standard dataset of images for testing registration nor a standard methodology or measures to compare results between methods. Moreover, most previous works in hyperspectral images registration validate the algorithms using datasets with transformations limited to a few scales and angles. Furthermore, few authors publish the executables or the source code of their algorithms, which makes it very difficult to compare methods. In this thesis, hyperspectral image registration methods have been developed and extensively tested on images of different types with a wide range of scales and angles (specifically, $4,680$ cases). The dataset contains real registration cases, as it includes pairs of images taken at different dates from different viewpoints and which present changes in infrastructure, buildings, crops, vegetation, and others. Therefore, one of the main contributions is the creation of a new dataset for testing. The proposed algorithms and the created dataset are available online and anyone can use them and follow the proposed testing methodology.

State-of-the-art registration methods for RGB and greyscale images have been exhaustively analysed in this thesis. Some of them have been implemented and adapted to efficiently exploit the spectral information in order to obtain better results and for a variety of registration parameters (scale, angle, and translation). The final design of the proposed algorithms, including the design decision on where to include the spectral information and how to exploit it, has been the result of a long experimentation process. This has made it possible to register images with scale factors of up to $12.5\times$, if we limit the results to the real case of pairs of images taken in different years.

Illumination changes due to atmospheric, seasonal or other conditions severely affect the registration. It has been observed that in some of the images taken by AVIRIS performing a band selection using classic image properties such as contrast or brightness improves the registration results in terms of number of scales that the method is able to register. However, this could not be extrapolated to all the test images and has therefore not been included or analysed in the thesis. Different methods available in the literature have also been proved to correct the pixel intensity differences as a pre-registration step, but the results were poor. The direct modification of band information, as we saw with PCA, affects the extraction and then the matching stages of feature-based methods. In this sense, a band selection method based on entropy and interband distance, which does not modify the original hyperspectral images, has been proposed. It was compared to state-of-the-art methods with better results and better computational performance.

Related to the previous points, different methods published in the literature for the registration of hyperspectral images have been analysed. As mentioned above, many of them ignore spectral information and only use one band of the images to register them. This thesis has confirmed that this can be efficient but not effective, as it fails to register images with large changes or scales. We have also shown that it is possible to create effective and efficient methods that efficiently run on parallel architectures. The first published methods for hyperspectral image registration on GPUs to the best of our knowledge are proposed in this thesis. As the results have shown, the registration problem will require one algorithm or another depending on different properties, such as image resolution or computation time. As a result, the user does not need a single method, but a toolbox with many methods to select the most suitable one for the particular problem.

In particular, the main contributions of this thesis and how the objectives have been met are summarised below:

1. *A Fourier-Mellin-based method for the registration of two hyperspectral images has been proposed.*

   The method, named HYFM, is based on principal component analysis, the multilayer fractional Fourier transform, a combination of log-polar maps, and peak processing. The Fourier transform is computed through the FRFT over a multilayer grid, which reduces the interpolation errors in the log-polar map. The FFT allows the efficient calculation of the correlation between two images.

   HYFM takes into account the multiband structure of the hyperspectral images through two strategies: the images are first processed using PCA, enabling dimensionality reduction and feature extraction. Second, the combination of log-polar maps is applied after the phase correlation, which allows each pair of components to contribute independently to the cross-correlation. This stage is aimed at highlighting some peaks in the log-polar map using information from different bands. A final peak processing stage is applied to discriminate the false peaks due to interpolation errors and rotationally dependent aliasing, as well as the images that do not match exactly.

   HYFM outperforms other area-based methods available in the literature by registering hyperspectral images with up to $7.5\times$ scalings.

2. *A feature-based registration pipeline for the efficient registration of hyperspectral images has been designed.*

   The proposed pipeline can be used with any feature detector and feature descriptor with few adaptations. It consists of six stages: band selection, feature detection, feature description, feature matching, band combination, and registration. It exploits the spectral information available in the hyperspectral images by performing feature detection and description in several bands of the images. For that purpose, we developed EBS, a band selection method focused on the hyperspectral image registration problem. It has been compared with other state-of-the-art methods showing better results.

   Furthermore, the registration pipeline incorporates spectral information into the descriptor. Thus, the features are described with spatial and spectral information making them more discriminating. Additionally, a higher number of false matches is discarded in the matching stage. Moreover, we have proposed an exhaustive search registration method based on histograms to estimate the transformation. It efficiently chooses the features

matches that will be finally used to compute the final geometric transformation. This proposal achieves better results than the most widely used method, RANSAC.

Three methods have been proposed following this pipeline: HSI–KAZE, HSI–SURF, and HSI–MSER. The results have shown that the proposed pipeline efficiently exploits the spectral information to improve registration, and it is robust and reusable. All of them were evaluated in detail in terms of the number of correct matches, correct match ratio, RMSE, or registration error.

HSI–KAZE is based on the detection of distinctive features called keypoints in a non-linear pyramidal scale space. Non-linear diffusion filtering allows applying a controlled blur while the important structures of the image are preserved. Thus, the robustness to scale invariance is improved. The method exploits the spectral information detecting keypoints in different selected bands and incorporating spectral information to the keypoint descriptors. Experiments show that HSI–KAZE is able to align images with scale factors up to $24\times$ by using spectral information.

HSI–SURF is a suitable solution when the problem is to register images with lower scale factors at low computational cost. The method is based on SURF as the keypoint detector and descriptor, and is specially adapted to hyperspectral images by incorporating the spectral information in different stages. It achieves correct registrations at scales of $11\times$ with lower RMSE and registration error.

HSI–MSER is based on the detection of distinctive regions by thresholding the hyperspectral images at different grey levels. An adaptation of the SIFT descriptor to regions is used as a spatial descriptor. The combination of this detector and descriptor with the exploitation of spectral information makes this method a very good registration algorithm, as it is able to align hyperspectral images with scales up to $15\times$. It is also computationally efficient due to its implementation following a tree structure.

3. *Efficient GPU implementations of HYFM and HSI–KAZE have been obtained by using CUDA*.

   We propose the first two published implementations of hyperspectral remote sensing registration algorithms on GPU to the extent of our knowledge: HYFM and HSI–KAZE. Different techniques and optimization strategies were applied to achieve efficient GPU computing.

The proposed implementations efficiently exploit the thousands of available threads on the GPU, obtaining considerable reductions in execution time as compared to CPU implementations. They were developed for running entirely on GPU, eliminating the costs of data transfers between the host and the device. Of the different optimisation techniques and strategies, the following stand out: the use of parallel reductions using warp-level primitives and warp-aggregated atomics, vectorial instructions, and efficient exploitation of the different memory spaces (shared memory, constant memory, texture memory, and global memory).

In the case of HSI–KAZE, experiments show a speedup of $13\times$ over an efficient multicore implementation in OpenMP, while for HYFM a speedup of $240\times$ is achieved compared to a sequential implementation in CPU.

Furthermore, both the implementations were integrated into a multi-GPU system in which each GPU operates separate datasets in parallel. The novelty of this proposal is that it carries out a two-level registration of multispectral images $10\times$ faster than a multicore OpenMP implementation. This demonstrates that the proposed methods can be effectively applied in multi-level registration.

4. *A dataset of hyperspectral images was created to test registration algorithms.*

   We have created a repository of hyperspectral images from different sensors and with pairs of images taken at different dates. The variety of characteristics of the images make this repository a standard dataset for testing registration algorithms. In addition to hyperspectral images, reference data and RGB images are published in the repository. Thus, any researcher can use them to evaluate their own method or to validate ours.

The proposed algorithms were evaluated for a wide variety of scale and rotation parameters, as well as compared in terms of registration accuracy to other methods in the literature. Nine hyperspectral images taken by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Reflective Optics System Imaging Spectrometer (ROSIS) sensors were used to evaluate the methods. They include urban or rural scenes and changes in different spatial structures and illumination. Furthermore, the algorithms were also tested in multispectral images captured by a MicaSense RedEdge MX sensor in different UAV flights. In all cases, our proposals overcome state-of-the-art methods.

The algorithms were developed to run in commodity hardware. For the sequential and OpenMP implementations, three CPUs were used: Intel Core i5-6600, Intel Core i7-4790, and

Intel Xeon E5-2623v4. For the GPU implementations, also three different GPUs corresponding to the Pascal and Turing architectures of NVIDIA were used: a GeForce GTX 1070, a Tesla P40, and a GeForce RTX 2070.

In view of the results presented in this thesis, we have achieved the main objectives proposed at the outset of this dissertation. We can conclude that the use of spectral information allows the alignment of multispectral and hyperspectral images with large and unknown scale factors and rotations. Additionally, the registration of multispectral and hyperspectral images can be performed in real-time by developing efficient and parallel algorithms for GPU.

10.1 Future work

In the conclusions, the work carried out has been presented, as well as the main challenges that this thesis solves. However, some problems were also mentioned that were beyond the scope of this thesis, but should be addressed in future work.

One important pending task in the remote sensing area is to build an user-friendly tool that allows us to access images in several databases at the same time, i.e. a database of databases. Thus, we can easily search for images of specific locations taken by AVIRIS, ROSIS, Hyperion, and Sentinel, among others. There are large datasets available in different public repositories, but they are often not used by the researchers interested in the registration problem because it is difficult to know if they contain images of a specific area of interest or because the repositories are not easily accessible. This tool could also be useful for applications such as change detection or image fusion, as it would provide images obtained by different sensors.

Another interesting task is the development of feature-based methods that are less sensitive to pixel intensity changes without losing good localisation accuracy, i.e. not at the cost of decreasing the number of correct matched features due to reduced indistinctiveness. This is a problem shared with many areas of remote sensing, for example, with change detection. Is this the same object or structure in both images? What can be considered change and what not? As explained before, sometimes registering images obtained in the same area is a challenge because the images are affected by changes due to the atmospheric conditions or seasonal variations, or because they are taken by different sensors.

Moreover, this thesis has effectively and efficiently solved the registration for hyperspectral images with a wide range of scales, angles, and translations. However, it has also addressed the problem of registration of images obtained by sensors on board UAVs, which has shown that

there are other problems associated with registration that require more sophisticated methods. These could be based on multilevel registration. A first level would involve a registration using the methods developed in this thesis. A second level would correct small misregistrations by computing e.g. elastic geometric transformations, since at low level the transformations needed to register the images with higher pixel accuracy are different for each part of the images. As a consequence, it is necessary to address the analysis of methods that search for transformations with more degrees of freedom or elastic registration techniques.

A few more directions that might be interesting to explore in the future are detailed below:

- *Coarse-to-fine registration method.* The methods developed in this thesis are good at dealing with large scales, rotations, and translations. However, hyperspectral remote sensing images usually present distortions that are not assumed by this model. HSI–KAZE, HSI–SURF or HSI–MSER can be used as the first stage of a multilevel registration method. These methods would perform a high-level registration that would be refined in a second registration stage to correct geometric deformations at a lower level. For the second level, non-rigid and elastic registration techniques would be studied.

- *Multimodal registration.* The increasing number and variety of hyperspectral sensors make it necessary to address this problem. In remote sensing image registration, changes in the images make the process difficult. For example, illumination changes because the atmospheric conditions were different when the images were taken or because they were taken in different seasons. In addition to these problems, in multimodal hyperspectral image registration, we have to cope with different spatial, spectral, and radiometric resolutions due to the use of different sensors.

- *Registration on distributed memory computer architectures.* To process large hyperspectral image datasets, it would be interesting to implement methods that can run on high performance computing systems such as distributed parallel systems, particularly if the images are of considerable size and have to be registered several times at different levels. One immediate option would be an extension of the proposed multi-GPU system that runs on a single shared memory machine (it has two CPUs that access the same memory space). Thus, a new level of parallelism would be added.

- *Apply computationally efficient versions of feature extraction algorithms to other problems*. Feature-based methods have traditionally been reviled for the excessive computational time they require, despite their distinctive and accurate feature extraction. The implementation of HSI–KAZE on GPU could be used for other problems thanks to its computational efficiency, such as content-based hyperspectral image retrieval.

Bibliography

[1] Álvaro Acción, Francisco Argüello, and Dora B. Heras. Dual-window superpixel data augmentation for hyperspectral image classification. *Applied Sciences*, 10(24), 2020.

[2] Andy Adinets. Optimized Filtering with Warp-Aggregated Atomics. `https://developer.nvidia.com/blog/cuda-pro-tip-optimized-filtering-warp-aggregated-atomics/`, 2021. [Online; accessed 15-06-2021].

[3] Motilal Agrawal, Kurt Konolige, and Morten Rufus Blas. Censure: Center surround extremas for realtime feature detection and matching. In *European Conference on Computer Vision*, pages 102–115. Springer, 2008.

[4] Naveed Akhtar, Faisal Shafait, and Ajmal Mian. Sparse spatio-spectral representation for hyperspectral image super-resolution. In *European conference on computer vision*, pages 63–78. Springer, 2014.

[5] David H Bailey and Paul N Swarztrauber. The fractional Fourier transform and applications. *SIAM review*, 33(3):389–404, 1991.

[6] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3):346–359, 2008.

[7] Jon A Benediktsson, Gabriele Cavallaro, Nicola Falco, Ihsen Hedhli, Vladimir A Krylov, Gabriele Moser, Sebastiano B Serpico, and Josiane Zerubia. Remote sensing data fusion: Markov models and mathematical morphology for multisensor, multiresolution, and multiscale image classification. In *Mathematical Models for Remote Sensing Image Processing*, pages 277–323. Springer, 2018.

[8] Peter Bunting, Frédéric Labrosse, and Richard Lucas. A multi-resolution area-based technique for automatic multi-modal image registration. *Image and Vision Computing*, 28(8):1203–1219, 2010.

[9] Stephen C Cain, Majeed M Hayat, and Ernest E Armstrong. Projection-based image registration in the presence of fixed-pattern noise. *IEEE transactions on image processing*, 10(12):1860–1872, 2001.

[10] Claude Cariou, Kacem Chehdi, and Steven Le Moan. BandClust: An unsupervised band reduction method for hyperspectral remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 8(3):565–569, 2011.

[11] Gabriele Cavallaro, Mauro Dalla Mura, Jón Atli Benediktsson, and Antonio Plaza. Remote sensing image classification using attribute filters defined over the tree of shapes. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7):3899–3911, 2016.

[12] Rohit Chandra, Leo Dagum, David Kohr, Ramesh Menon, Dror Maydan, and Jeff McDonald. *Parallel programming in OpenMP*. Morgan kaufmann, 2001.

[13] Chein-I Chang. *Hyperspectral imaging: techniques for spectral detection and classification*, volume 1. Springer Science & Business Media, 2003.

[14] Subhasis Chaudhuri and Ketan Kotwal. *Hyperspectral image fusion*. Springer, 2013.

[15] H-M Chen, Manoj K Arora, and Pramod K Varshney. Mutual information-based image registration for remote sensing data. *International Journal of Remote Sensing*, 24(18):3701–3706, 2003.

[16] Qin-sheng Chen, Michel Defrise, and Frank Deconinck. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 16(12):1156–1168, 1994.

[17] Shu-Jie Chen, Hui-Liang Shen, Chunguang Li, and John H Xin. Normalized total gradient: a new measure for multispectral image registration. *IEEE Transactions on Image Processing*, 27(3):1297–1310, 2018.

[18] Siyue Chen, Qing Guo, Henry Leung, and Eloi Bosse. A maximum likelihood approach to joint image registration and fusion. *IEEE Transactions on Image Processing*, 20(5):1363–1372, 2011.

[19] Nigel Chou, Joseph A Izatt, and Sina Farsiu. Generalized pseudo-polar Fourier grids and applications in registering ophthalmic optical coherence tomography images. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 807–811. IEEE, 2009.

[20] Copernicus Open Access Hub . Complete, free and open access to Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P images. `https://scihub.copernicus.eu/userguide/`. [Online; accessed 02-06-2021].

[21] Xiaolong Dai and Siamak Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(5):1566–1577, 1998.

[22] Majolie Djokam, Maxleene Sandasi, Weiyang Chen, Alvaro Viljoen, and Ilze Vermaak. Hyperspectral imaging as a rapid quality control method for herbal tea blends. *Applied Sciences*, 7(3):268, 2017.

[23] Pinliang Dong and Qi Chen. *LiDAR remote sensing and applications*. CRC Press, 2017.

[24] Leidy P Dorado-Muñoz, Miguel Velez-Reyes, Amit Mukherjee, and Badrinath Roysam. A vector SIFT detector for interest point detection in hyperspectral imagery. *IEEE transactions on Geoscience and Remote sensing*, 50(11):4521–4533, 2012.

[25] Earth Resources Observation and Science (EROS) Center. Earth Observing One (EO-1) - Hyperion. `https://www.usgs.gov/centers/eros/science/usgs-eros-archive-earth-observing-one-eo-1-hyperion`. [Online; accessed 02-06-2021].

[26] Charles Elachi and Jakob J Van Zyl. *Introduction to the physics and techniques of remote sensing*. John Wiley & Sons, 2021.

[27] Hector Erives and Glenn J Fitzgerald. Automated registration of hyperspectral images for precision agriculture. *Computers and electronics in agriculture*, 47(2):103–119, 2005.

[28] Hector Erives and Glenn J Fitzgerald. Automatic subpixel registration for a tunable hyperspectral imaging system. *IEEE Geoscience and remote sensing letters*, 3(3):397–400, 2006.

[29] Adrian N Evans and Xin U Liu. A morphological gradient approach to color edge detection. *IEEE Transactions on Image Processing*, 15(6):1454–1463, 2006.

[30] Pablo F. Alcantarilla, Jesús Nuevo, and Adrien Bartoli. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell*, 34(7):1281–1298, 2011.

[31] Nicola Falco, Jón Atli Benediktsson, and Lorenzo Bruzzone. Spectral and spatial classification of hyperspectral images based on ica and reduced morphological attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11):6223–6240, 2015.

[32] Nicola Falco, Gabriele Cavallaro, Prashanth R. Marpu, and Jon Atli Benediktsson. Unsupervised change detection analysis to multi-channel scenario based on morphological contextual analysis. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3374–3377, 2016.

[33] Zhe Fan, Christoph Vetter, Christoph Guetter, Daphne Yu, Rüdiger Westermann, Arie Kaufman, and Chenyang Xu. Optimized GPU implementation of learning-based non-rigid multi-modal registration. In *Medical Imaging*, pages 69142Y–69142Y. International Society for Optics and Photonics, 2008.

[34] Mathieu Fauvel, Yuliya Tarabalka, Jon Atli Benediktsson, Jocelyn Chanussot, and James C Tilton. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3):652–675, 2012.

[35] Yao-Ze Feng and Da-Wen Sun. Application of hyperspectral imaging in food safety inspection and control: a review. *Critical reviews in food science and nutrition*, 52(11):1039–1058, 2012.

[36] Pablo Fernández Alcantarilla, Adrien Bartoli, and Andrew J Davison. KAZE features. In *European Conference on Computer Vision*, pages 214–227. Springer, 2012.

[37] Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. Towards a multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. In J. Vigo Aguiar, editor, *Computational and Mathematical Methods in Science and Engineering*, 2018.

[38] Jorge Fernández-Fabeiro, Álvaro Ordóñez, Arturo González-Escribano, and Dora B. Heras. A multi-device version of the HYFMGPU algorithm for hyperspectral scenes registration. *Journal of Supercomputing*, 75(3):1551–1564, 2019.

[39] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[40] Hernan Flores, Sandra Lorenz, Robert Jackisch, Laura Tusa, I. Cecilia Contreras, Robert Zimmermann, and Richard Gloaguen. Uas-based hyperspectral environmental monitoring of acid mine drainage affected waters. *Minerals*, 11(2), 2021.

[41] Africa I Flores-Anderson, Robert Griffin, Margaret Dix, Claudia S Romero-Oliva, Gerson Ochaeta, Juan Skinner-Alvarado, Maria Violeta Ramirez Moran, Betzy Hernandez, Emil Cherrington, Benjamin Page, et al. Hyperspectral satellite remote sensing of water quality in lake atitlán, guatemala. *Frontiers in Environmental Science*, 8:7, 2020.

[42] Open Source Vision Foundation. Open Source Computer Vision Library (OpenCV). `https://opencv.org/`, 2021. [Online; accessed 27-05-2021].

[43] Python Software Foundation. Python Official Page. `https://www.python.org`, 2021. [Online; accessed 27-05-2021].

[44] Yuanyuan Fu, Guijun Yang, Zhenhai Li, Heli Li, Zhenhong Li, Xingang Xu, Xiaoyu Song, Yunhe Zhang, Dandan Duan, Chunjiang Zhao, et al. Progress of hyperspectral data processing and modelling for cereal crop nitrogen monitoring. *Computers and Electronics in Agriculture*, 172:105321, 2020.

[45] Pedro G. Bascoy, Pablo Quesada-Barriuso, Dora B. Heras, Francisco Argüello, Begüm Demir, and Lorenzo Bruzzone. Extended attribute profiles on GPU applied to hyperspectral image classification. *The Journal of Supercomputing*, 75(3):1565–1579, 2019.

[46] Pedro G. Bascoy, Alberto S. Garea, Dora B. Heras, Francisco Argüello, and Álvaro Ordóñe. Texture-based analysis of hydrographical basins with multispectral imagery. In Christopher M. U. Neale and Antonino Maltese, editors, *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI*, volume 11149, pages 225 – 234. International Society for Optics and Photonics, SPIE, 2019.

[47] Francisco J Galdames, Claudio A Perez, Pablo A Estévez, and Martin Adams. Rock lithological classification by hyperspectral, range 3d and color images. *Chemometrics and Intelligent Laboratory Systems*, 189:138–148, 2019.

[48] Lianru Gao, Dan Yao, Qingting Li, Lina Zhuang, Bing Zhang, and José M Bioucas-Dias. A new low-rank representation based hyperspectral image denoising method for mineral mapping. *Remote Sensing*, 9(11):1145, 2017.

[49] Pedram Ghamisi, Javier Plaza, Yushi Chen, Jun Li, and Antonio J Plaza. Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geoscience and Remote Sensing Magazine*, 5(1):8–32, 2017.

[50] GNU. GCC, the GNU Compiler Collection. `https://gcc.gnu.org`, 2021. [Online; accessed 27-05-2021].

[51] Hernani Goncalves, Luís Corte-Real, and Jose A Goncalves. Automatic image registration through image segmentation and SIFT. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7):2589–2600, 2011.

[52] Maoguo Gong, Shengmeng Zhao, Licheng Jiao, Dayong Tian, and Shuang Wang. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7):4328–4338, 2014.

[53] A. Ardeshir Goshtasby. *Image registration: Principles, tools and methods*. Springer Science & Business Media, 2012.

[54] Aoife A Gowen, Colm P O'Donnell, Patrick J Cullen, Gérard Downey, and Jesus M Frias. Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12):590–598, 2007.

[55] Grupo de Inteligencia Computacional de la Universidad del País Vasco (UPV/EHU). Hyperspectral Remote Sensing Scenes. http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes, 2020. [Online; accessed 14-05-2021].

[56] Babak Haghighi, Nathan D Ellingwood, Youbing Yin, Eric A Hoffman, and Ching-Long Lin. A GPU-based symmetric non-rigid image registration method in human lung. *Medical & biological engineering & computing*, 56(3):355–371, 2018.

[57] Fredric J Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, 66(1):51–83, 1978.

[58] Mahmudul Hasan, Xiuping Jia, Antonio Robles-Kelly, Jun Zhou, and Mark R Pickering. Multi-spectral remote sensing image registration via spatial relationship analysis on SIFT keypoints. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pages 1011–1014. IEEE, 2010.

[59] Fenghua Huang, Ying Yu, and Tinghao Feng. Hyperspectral remote sensing image change detection based on tensor and deep learning. *Journal of Visual Communication and Image Representation*, 58:233–244, 2019.

[60] Rey Otero Ives and Mauricio Delbracio. Anatomy of the SIFT method. *Image Processing On Line*, 4:370–396, 2014.

[61] Jet Propulsion Laboratory (NASA). AVIRIS database. https://aviris.jpl.nasa.gov/data/index.html. [Online; accessed 20-05-2021].

[62] Shun'ichi Kaneko, Yutaka Satoh, and Satoru Igarashi. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36(5):1165–1173, 2003.

[63] Xudong Kang, Shutao Li, and Jón Atli Benediktsson. Feature extraction of hyperspectral images with image fusion and recursive filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6):3742–3752, 2013.

[64] Yosi Keller, Amir Averbuch, and Moshe Israeli. Pseudopolar-based estimation of large translations, rotations, and scalings in images. *Image Processing, IEEE Transactions on*, 14(1):12–22, 2005.

[65] Jeffrey P Kern and Marios S Pattichis. Robust multispectral image registration using mutual-information models. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1494–1505, 2007.

[66] Marek Kovar, Marian Brestic, Oksana Sytar, Viliam Barek, Pavol Hauptvogel, and Marek Zivcak. Evaluation of hyperspectral reflectance parameters to assess the leaf water content in soybean. *Water*, 11(3):443, 2019.

[67] Kavitha Kuppala, Sandhya Banda, and Thirumala Rao Barige. An overview of deep learning methods for image registration with focus on feature-based approaches. *International Journal of Image and Data Fusion*, 11(2):113–135, 2020.

[68] Chiman Kwan, Joon Hee Choi, Stanley H Chan, Jin Zhou, and Bence Budavari. A super-resolution and fusion approach to enhancing hyperspectral images. *Remote Sensing*, 10(9):1416, 2018.

[69] Charis Lanaras, Emmanuel Baltsavias, and Konrad Schindler. Hyperspectral super-resolution by coupled spectral unmixing. In *Proceedings of the IEEE international conference on computer vision*, pages 3586–3594, 2015.

[70] Eve Laroche-Pinel, Mohanad Albughdadi, Sylvie Duthoit, Véronique Chéret, Jacques Rousseau, and Harold Clenet. Understanding vine hyperspectral signature through different irrigation plans: A first step to monitor vineyard water status. *Remote Sensing*, 13(3):536, 2021.

[71] Jacqueline Le Moigne, Nathan S Netanyahu, and Roger D Eastman. *Image registration for remote sensing*. Cambridge University Press, 2011.

[72] Liangzhi Li, Ling Han, Mingtao Ding, Zhiheng Liu, and Hongye Cao. Remote sensing image registration based on deep learning regression model. *IEEE Geoscience and Remote Sensing Letters*, 2020.

[73] Qiaoliang Li, Suwen Qi, Yuanyuan Shen, Dong Ni, Huisheng Zhang, and Tianfu Wang. Multispectral image alignment with nonlinear scale-invariant keypoint and enhanced local feature matrix. *IEEE Geoscience and Remote Sensing Letters*, 12(7):1551–1555, 2015.

[74] Qiaoliang Li, Guoyou Wang, Jianguo Liu, and Shaobo Chen. Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 6(2):287–291, 2009.

[75] Yuan Lin and Vinod Grover. Using CUDA Warp-Level Primitives. `https://developer.nvidia.com/blog/using-cuda-warp-level-primitives/`, 2021. [Online; accessed 15-06-2021].

[76] Hanzhou Liu, Baolong Guo, and Zongzhe Feng. Pseudo-log-polar Fourier transform for image registration. *Signal Processing Letters, IEEE*, 13(1):17–20, 2006.

[77] Sicong Liu, Daniele Marinelli, Lorenzo Bruzzone, and Francesca Bovolo. A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges. *IEEE Geoscience and Remote Sensing Magazine*, 7(2):140–158, 2019.

[78] Javier López Fandiño. *Efficient multitemporal change detection techniques for hyperspectral images on GPU*. PhD thesis, Universidade de Santiago de Compostela, 2018.

[79] Javier López-Fandiño, Alberto S Garea, Dora B. Heras, and Francisco Argüello. Stacked autoencoders for multiclass change detection in hyperspectral images. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1906–1909. IEEE, 2018.

[80] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[81] Wenping Ma, Zelian Wen, Yue Wu, Licheng Jiao, Maoguo Gong, Yafei Zheng, and Liang Liu. Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, 14(1):3–7, 2017.

[82] Hani Mahdi and Aly A Farag. Image registration in multispectral data sets. In *Proceedings. International Conference on Image Processing*, volume 2, pages II–II. IEEE, 2002.

[83] Adolfo Martínez-Usó, Filiberto Pla, José M. Sotoca, and Pedro García-Sevilla. Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience & Remote Sensing*, 45(12):4158–4171, December 2007.

[84] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767, 2004.

[85] F Mehmet, Yasemin Yardimci, Alptekin Temlzel, et al. Registration of multispectral satellite images with orientation-restricted SIFT. In *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, volume 3, pages III–243. IEEE, 2009.

[86] Krystian Mikolajczyk, Tinne Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky, Timor Kadir, and Luc Van Gool. A comparison of affine region detectors. *International journal of computer vision*, 65(1-2):43–72, 2005.

[87] Michael Moeller, Todd Wittman, and Andrea L Bertozzi. A variational approach to hyperspectral image fusion. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, volume 7334, page 73341E. International Society for Optics and Photonics, 2009.

[88] Amit Mukherjee, Miguel Velez-Reyes, and Badrinath Roysam. Interest points for hyperspectral image data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3):748–760, 2009.

[89] Pynar Muyan-Ozcelik, John D Owens, Junyi Xia, and Sanjiv S Samant. Fast deformable registration on the GPU: A CUDA implementation of demons. In *Computational Sciences and Its Applications, 2008. ICCSA'08. International Conference on*, pages 223–233. IEEE, 2008.

[90] Koushik Nagasubramanian, Sarah Jones, Soumik Sarkar, Asheesh K Singh, Arti Singh, and Baskar Ganapathysubramanian. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant methods*, 14(1):1–13, 2018.

[91] Laurent Najman and Michel Couprie. Building the component tree in quasi-linear time. *IEEE Transactions on image processing*, 15(11):3531–3539, 2006.

[92] NVIDIA. Whitepaper NVIDIA GeForce GTX 1080. `http://international.download.nvidia.com/geforce-com/international/pdfs/`

GeForce_GTX_1080_Whitepaper_FINAL.pdf. [Online; accessed 25-05-2021].

[93] NVIDIA. Whitepaper NVIDIA Turing GPU Architecture. https://images.nvidia.com/aem-dam/en-zz/Solutions/design-visualization/technologies/turing-architecture/NVIDIA-Turing-Architecture-Whitepaper.pdf. [Online; accessed 17-06-2021].

[94] NVIDIA. NVIDIA's Next Generation CUDA Compute Architecture: Kepler GK110/210. https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/tesla-product-literature/NVIDIA-Kepler-GK110-GK210-Architecture-Whitepaper.pdf, 2014. [Online; accessed 26-05-2021].

[95] NVIDIA. Whitepaper: NVIDIA Tesla P100. https://images.nvidia.com/content/pdf/tesla/whitepaper/pascal-architecture-whitepaper.pdf, 2016. [Online; accessed 26-05-2021].

[96] NVIDIA. CUB Library 1.8.0. https://nvlabs.github.io/cub/, 2018. [Online; accessed 27-05-2021].

[97] NVIDIA. cuBLAS Library User's Guide. http://docs.nvidia.com/cuda/pdf/CUBLAS_Library.pdf, May 2021. [Online; accessed 27-05-2021].

[98] NVIDIA. CUDA C++ Best Practices Guide. https://docs.nvidia.com/cuda/cuda-c-best-practices-guide/index.html, 2021. [Online; accessed 22-05-2021].

[99] NVIDIA. CUDA C++ Programming Guide. https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html, 2021. [Online; accessed 24-05-2021].

[100] NVIDIA. CUDA Toolkit. https://developer.nvidia.com/cuda-toolkit, 2021. [Online; accessed 27-05-2021].

[101] NVIDIA. cuFFT Library User's Guide. http://docs.nvidia.com/cuda/pdf/CUFFT_Library.pdf, May 2021. [Online; accessed 27-05-2021].

[102] NVIDIA. cuSOLVER Library User's Guide. http://docs.nvidia.com/cuda/pdf/CUSOLVER_Library.pdf, May 2021. [Online; accessed 27-05-2021].

[103] NVIDIA. NVIDIA Performance Primitives (NPP) v11.3.0 User's Guide. https://docs.nvidia.com/cuda/npp/index.html, 2021. [Online; accessed 27-05-2021].

[104] NVIDIA. Thrust Quick Start Guide. http://docs.nvidia.com/cuda/pdf/Thrust_Quick_Start_Guide.pdf, May 2021. [Online; accessed 27-05-2021].

[105] OpenMP Architecture Review Board. OpenMP Application Programming Interface version 5.1. https://www.openmp.org/wp-content/uploads/OpenMP-API-Specification-5-1.pdf, November 2020. [Online; accessed 24-05-2021].

[106] Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Fourier-Mellin registration of two hyperspectral images. *International Journal of Remote Sensing*, 38(11):3253–3273, 2017.

[107] Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. GPU accelerated FFT-Based registration of hyperspectral scenes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4869–4878, 2017.

[108] Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Transformada de Fourier aplicada al alineamiento de imágenes multidimensionales en GPU. In Rafael Asenjo, Ángeles Navarro, Arturo González-Escribano, Diego R. Llanos, Sergio Cuenca Asensi, and Jesús González Peñalver, editors, *Jornadas SARTECO*, 2017.

[109] Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Alignment of hyperspectral images using KAZE features. *Remote Sensing*, 10(5), 2018.

[110] Álvaro Ordóñez, Francisco Argüello, Dora B. Heras, and Begüm Demir. GPU-Accelerated registration of hyperspectral images using KAZE features. *Journal of Supercomputing*, 76(12):9478–9492, 2020.

[111] Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Exploring the registration of remote sensing images using HSI-KAZE in graphical units. In J. Vigo Aguiar, editor, *Computational and Mathematical Methods in Science and Engineering*, 2019.

[112] Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. SURF-based registration for hyperspectral images. In *International Geoscience and Remote Sensing Symposium*, pages 63–66. IEEE, 2019.

[113] Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Exploring the MSER-based hyperspectral remote sensing image registration. In *Image and Signal Processing for Remote Sensing XXVI*. SPIE, 2020.

[114] Álvaro Ordóñez, Dora B. Heras, and Francisco Argüello. Comparing area-based and feature-based methods for co-registration of multispectral bands on GPU. In *International Geoscience and Remote Sensing Symposium*, pages 1575–1578. IEEE, 2021.

[115] Álvaro Ordóñez, Francisco Argüello, and Dora B. Heras. Repository of hyperspectral images for image registration. `https://gitlab.citius.usc.es/hiperespectral/RegistrationRepository`, 2021. [Online; accessed 14-05-2021].

[116] W. Pan, K. Qin, and Y. Chen. An adaptable-multilayer fractional Fourier transform approach for image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(3):400–414, March 2009.

[117] Mercedes E Paoletti, Juan M Haut, Xuanwen Tao, Javier Plaza Miguel, and Antonio Plaza. A new GPU implementation of support vector machines for fast hyperspectral image classification. *Remote Sensing*, 12(8):1257, 2020.

[118] Yongeun Park, JongCheol Pyo, Yong Sung Kwon, YoonKyung Cha, Hyuk Lee, Taegu Kang, and Kyung Hwa Cho. Evaluating physico-chemical influences on cyanobacterial blooms using hyperspectral images in inland water, korea. *Water research*, 126:319–328, 2017.

[119] Antonio Plaza, Javier Plaza, and Hugo Vegas. Improving the performance of hyperspectral image and signal processing algorithms using parallel, distributed and specialized hardware-based systems. *Journal of Signal Processing Systems*, 61(3):293–315, 2010.

[120] Yinghui Quan, Xian Zhong, Wei Feng, Gabriel Dauphin, Lianru Gao, and Mengdao Xing. A novel feature extension method for the forest disaster monitoring using multispectral data. *Remote Sensing*, 12(14):2261, 2020.

[121] Pablo Quesada-Barriuso, Dora Blanco Heras, and Francisco Argüello. GPU accelerated waterpixel algorithm for superpixel segmentation of hyperspectral images. *The Journal of Supercomputing*, pages 1–13, 2021.

[122] B. Srinivasa Reddy and Biswanath N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996.

[123] John Alan Richards. *Remote sensing digital image analysis*, volume 3. Springer, 1999.

[124] Edward Rosten and Tom Drummond. Machine learning for high-speed corner detection. In *European conference on computer vision*, pages 430–443. Springer, 2006.

[125] Neven G Rostom, Adel A Shalaby, Yousry M Issa, and Ahmed A Afifi. Evaluation of mariut lake water quality using hyperspectral remote sensing and laboratory works. *The Egyptian Journal of Remote Sensing and Space Science*, 20:S39–S48, 2017.

[126] Sudhakar Sah, Jan Vanek, YoungJun Roh, and Ratul Wasnik. GPU accelerated real time rotation, scale and translation invariant image registration method. In *Image Analysis and Recognition*, pages 224–233. Springer, 2012.

[127] Jason Sanders and Edward Kandrot. *CUDA by example: an introduction to general-purpose GPU programming*. Addison-Wesley Professional, 2010.

[128] Silvia Serranti, Aldo Gargiulo, and Giuseppe Bonifazi. Hyperspectral imaging for process and quality control in recycling plants of polyolefin flakes. *Journal of Near Infrared Spectroscopy*, 20(5):573–581, 2012.

[129] Ramtin Shams, Parastoo Sadeghi, Rodney Kennedy, and Richard Hartley. Parallel computation of mutual information on the GPU with application to real-time registration of 3D medical images. *Computer methods and programs in biomedicine*, 99(2):133–146, 2010.

[130] Susan C Steele-Dunne, Heather McNairn, Alejandro Monsivais-Huertero, Jasmeet Judge, Pang-Wei Liu, and Kostas Papathanassiou. Radar remote sensing of agricultural

canopies: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5):2249–2273, 2017.

[131] Hao Sun, Xiangtao Zheng, and Xiaoqiang Lu. A supervised segmentation network for hyperspectral image classification. *IEEE Transactions on Image Processing*, 30:2810–2825, 2021.

[132] Mustafa Teke and Alptekin Temizel. Multi-spectral satellite image registration using scale-restricted SURF. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2310–2313. IEEE, 2010.

[133] FJA Van Ruitenbeek, HMA van der Werff, WH Bakker, FD van der Meer, and KAA Hein. Measuring rock microstructure in hyperspectral mineral maps. *Remote sensing of environment*, 220:94–109, 2019.

[134] Vasily Volkov. Better performance at lower occupancy. In *Proceedings of the GPU technology conference, GTC*, volume 10, page 16. San Jose, CA, 2010.

[135] Jun Wang, Zhiyong Xu, and Jianlin Zhang. Image registration with hyperspectral data based on Fourier-Mellin transform. *Int. J. Signal. Process. Syst*, 1:107–110, 2013.

[136] Shuang Wang, Dou Quan, Xuefeng Liang, Mengdan Ning, Yanhe Guo, and Licheng Jiao. A deep learning framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145:148–164, 2018. Deep Learning RS Data.

[137] Joe H Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.

[138] Marc Wieland and Sandro Martinis. A modular processing chain for automated flood monitoring from multi-spectral satellite data. *Remote Sensing*, 11(19):2330, 2019.

[139] Xian-xiang Wu, Bao-long Guo, and Juan Wang. Octa-log-polar Fourier transform for image registration. In *Information Assurance and Security, 2009. IAS'09. Fifth International Conference on*, volume 1, pages 601–604. IEEE, 2009.

[140] Yue Wu, Qiguang Miao, Wenping Ma, Maoguo Gong, and Shanfeng Wang. Psosac: Particle swarm optimization sample consensus algorithm for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 2017.

[141] Zebin Wu, Linlin Shi, Jun Li, Qicong Wang, Le Sun, Zhihui Wei, Javier Plaza, and Antonio Plaza. GPU parallel implementation of spatially adaptive hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(4):1131–1143, 2017.

[142] Famao Ye, Yanfei Su, Hui Xiao, Xuqing Zhao, and Weidong Min. Remote sensing image registration using convolutional neural network features. *IEEE Geoscience and Remote Sensing Letters*, 15(2):232–236, 2018.

[143] Zheng Yi, Cao Zhiguo, and Xiao Yang. Multi-spectral remote image registration based on SIFT. *Electronics Letters*, 44(2):107–108, 2008.

[144] Naoto Yokoya, Jonathan Cheung-Wai Chan, and Karl Segl. Potential of resolution-enhanced hyperspectral data for mineral mapping using simulated enmap and sentinel-2 images. *Remote Sensing*, 8(3):172, 2016.

[145] Orestis Zachariadis, Andrea Teatini, Nitin Satpute, Juan Gómez-Luna, Onur Mutlu, Ole Jakob Elle, and Joaquín Olivares. Accelerating b-spline interpolation on GPUs: Application to medical image registration. *Computer methods and programs in biomedicine*, 193:105431, 2020.

[146] Huisheng Zhang, Xinyu Liu, Ling Li, Jie Rao, Qiaoliang Li, Siping Chen, and Tianfu Wang. Accurate multi-spectral image registration based on scale invariant feature. In *Computer Science and Network Technology (ICCSNT), 2012 2nd International Conference on*, pages 847–852. IEEE, 2012.

[147] Qian Zhang, Zhiguo Cao, Zhongwen Hu, Yonghong Jia, and Xiaoliang Wu. Joint image registration and fusion for panchromatic and multispectral images. *IEEE geoscience and remote sensing letters*, 12(3):467–471, 2015.

[148] Yunsheng Zhang, Peilong Zhou, Yue Ren, and Zhengrong Zou. GPU-accelerated large-size VHR images registration via coarse-to-fine matching. *Computers & Geosciences*, 66:54–65, 2014.

[149] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

List of Figures

List of Tables