3 Estado da arte em classificação de imagens de alta resolução

Com a recente disponibilidade de imagens de alta resolução produzidas por sensores orbitais como IKONOS e QUICKBIRD se tornou-se possível modelar com maior nível de detalhes características da cobertura do solo do que era possível com imagens de média ou baixa resolução. Imagens IKONOS, por exemplo, têm pixels de 1m por 1m, no caso da banda pancromática. Essa resolução é o suficiente para visualizar com detalhes, por exemplo, o teto de uma casa, árvores, jardins e ruas.

Graças à elevada resolução espacial é possível consolidar a informação contida em imagens de alta resolução com dados de um SIG que normalmente se encontraram disponíveis em escalas grandes, principalmente trabalhando a níveis locais.

Nesse contexto uma série de trabalhos têm sido desenvolvidos utilizando diferentes técnicas automáticas de representação e interpretação em imagens de alta resolução. Essas técnicas envolvem não somente diferentes modelos de tratamento de dados, mas também englobam diferentes métodos de representação da informação relevante para a tarefa de interpretação.

Esse capítulo aborda o estado da arte em classificação de imagens de alta resolução. O capítulo inicia com métodos baseados em pixel, e passa depois para uma abordagem mais generalista e mais moderna baseada em objetos ou segmentos. O procedimento de segmentação e as diferentes formas de representação de conhecimento são também considerados neste capítulo. Por último uma descrição dos métodos atuais de representação de conhecimento multitemporal é apresentada.

3.1 Métodos baseados em pixels

Métodos automáticos de classificação baseados em pixels foram amplamente utilizados em procedimentos de classificação em média e baixa resolução e têm sido utilizados ainda em imagens de alta resolução. A tarefa de classificação de imagens de alta resolução pode ser considerada difícil, tendo em vista a elevada complexidade e

composição dos tipos de cobertura do solo que podem ser discriminados nesse tipo de imagem, como se demonstra, por exemplo em [38].

Várias abordagens foram realizadas utilizando apenas dados espectrais e classificadores como o de máxima probabilidade e distância euclidiana [39]. Em [40, 41, 42, 56 e 57] diversos métodos de classificação baseados em lógica nebulosa, foram utilizados para a classificação de imagens de alta resolução. Tais métodos baseados somente na informação espectral alcançam resultados modestos, quando aplicados a imagens de alta resolução [42].

O conhecimento estrutural e contextual foi considerado em [43] utilizando uma análise de vizinhança dos pixel. A informação numa vizinhança em torno de cada pixel foi usada para extrair dados sobre o relacionamento espacial entre os pixels, gerando desta forma um atributo de contexto. Esse método foi avaliado utilizando imagens IKONOS pancromáticas e fotos aéreas em conjunto com um classificador baseado em redes neurais. Os resultados demonstraram claramente os benefícios do uso dos atributos de contexto.

Um novo método multi-escala foi proposto em [44] que combina imagens IKONOS pancromáticas e multiespectrais com fotografias aéreas. Neste trabalho o conjunto de dados é então tratado por classificadores de máxima probabilidade e classificadores baseados em árvore de regressão (CART). O resultado desse trabalho deu uma medida objetiva da importância da resolução espacial na tarefa de interpretação. O trabalho mostra ainda que a informação multi-escala pode ser utilizada com uma espécie de contexto no processo de classificação. Foi mostrado, por exemplo, que imagens com 1m de resolução favorecem a interpretação de alguns tipos de classes, por exemplo, sombreamento entre as árvores

A utilização da informação de textura na classificação de imagens de sensores remotos também vem sendo bem explorada desde há muito tempo.

Um [45] novo método de representação da textura foi proposto baseado no conceito de visões de um determinado padrão de textura. Basicamente o conceito aplicado é que uma determinada classe possui inúmeros padrões de textura e ainda que cada um desses padrões possui algumas visões. Essas visões foram obtidas com dados de treinamento e através de uma série de parâmetros de textura. Foi desenvolvido então um algoritmo de classificação semelhante ao KNN, o *K-View*. Esse método em linhas gerais atribui o padrão analisado a classe que possui “visões” mais semelhantes. Esses métodos de extração de textura e de classificação foram aplicados a imagens de satélite

de alta resolução resultando num incremento significativo da capacidade de discriminação de classes de cobertura de solo.

Em [46] análise de textura foi aplicada em imagens IKONOS e QUICKBIRD para mapeamento de árvores, pomares e plantações de café em Uganda. Foram utilizados parâmetros de textura de Haralick, calculados com base na matriz de coocorrência de níveis de cinza ou GLCM (*gray-level cooccurrence matrix*). Os resultados relatados no trabalho foram muito positivos principalmente no que disse respeito ao reconhecimento de árvores e pomares. Porém o autor relata uma dificuldade para determinar a dimensão ideal da vizinhança em torno de cada pixel a ser usada no cálculo dos atributos de textura. Um outro trabalho [47] reafirma a importância da informação de textura no processo de classificação, e ressalta igualmente a necessidade de uma escolha adequada do tamanho dessas janelas de análise.

Além dos estudos sobre atributos muito se tem investigado sobre diferentes modelos de classificadores para as aplicações em sensoriamento remoto.

Uma comparação entre métodos estatísticos e redes neurais aplicados diretamente em classificação de imagens de sensores remotos é descrita em [48] utilizando múltiplas fontes de dados LANDSAT MMS e dados topográficos. Os experimentos mostraram que os modelos baseados em redes neurais têm taxas de reconhecimento potencialmente superiores aos métodos estatísticos convencionais, mas perdem em termos de desempenho computacional. Esse baixo desempenho se justifica pelo elevado número de padrões que se tem em imagens de alta resolução, o que onera principalmente a fase de treinamento da rede neural. Os experimentos também mostraram a importância de se condicionar os dados adequadamente quando se utilizam redes neurais. Pequenas mudanças no condicionamento dos dados podem alterar sensivelmente o resultado da classificação.

Em [49] foi utilizado um método estatístico multivariado baseado em SVD (*Singular Value Decomposition*) e análise de vetores chaves (*key vector analysis*) oferecendo uma solução alternativa a redes neurais artificiais para a classificação de imagens aéreas de alta resolução. A justificativa dada pelos autores para a procura de um método alternativo as redes neurais foi mais uma vez o baixo desempenho computacional das redes neurais, quase sempre devido à fase de treinamento, e a alta sensibilidade das redes neurais aos seus pesos iniciais, que são gerados de forma randômica, na maioria dos casos.

Imagens de alta resolução espectral, obtidas a partir de sensor aerotransportado, AVIRIS, de 1991 foram analisadas em [50]. Esse tipo de imagem possui 224 canais, ou bandas espectrais, o que impossibilita, ou pelo menos torna difícil, uma análise estatística multivariada. A solução adotada pelo autor foi reduzir a dimensionalidade dos dados e aplicar em seguida um classificador não paramétrico. Foram utilizadas, mais uma vez, redes neurais artificiais como métodos não paramétricos de classificação. Para a redução de dimensionalidade foi utilizado um método chamado de DBFE (*Decision Boundary Feature Extraction*) [51][52][53]. A dimensionalidade foi reduzida de 224 para 35 e 36 bandas em dois experimentos realizados. Vários classificadores estatísticos foram comparados com as redes neurais artificiais para avaliar como cada um se comportava diante da redução de dimensionalidade dos dados. Classificadores baseados em redes neurais se mostraram superiores em termos de reconhecimento global. Os experimentos realizados mostraram ainda uma “singular” dificuldade do algoritmo de máxima verossimilhança nos casos em que determinadas classes possuem poucos padrões de treinamento, o que impossibilita uma boa estimação da matriz de covariância.

3.2 Classificação baseada em objetos

Os métodos de interpretação e classificação baseados em pixels têm se mostrados insatisfatórios para imagens de alta resolução. Imagens de alta resolução classificadas por métodos baseados em pixel tendem a mostrar pixels isolados como uma classe claramente inserida em uma área que representa uma outra classe [54].

Por outro lado surge a abordagem baseada em objetos que supõe uma pré-organização dos pixels em grupos, denominados objetos ou segmentos. O processo de extração ou formulação de grupos ou segmentos é conhecido como segmentação. A análise por segmentos pode realmente simplificar a complexidade de informação contida em um pixel de uma imagem, especialmente em imagens de alta resolução, tanto espacial quanto radiométrica [55].

Do ponto de vista da arquitetura do classificador a classificação de segmentos é idêntica à classificação baseada em pixels. A diferença principal é que o padrão é um segmento e não um pixel. Por esse motivo a maioria das técnicas para a classificação baseada em pixels, em princípio, funciona bem para a classificação de objetos. A maior

diferença está nos atributos ou no conhecimento que se pode expressar utilizando uma abordagem por segmentos.

Com o objetivo de investigar novas metodologias de classificação explorando dados de SIG para a cidade do Rio de Janeiro, Rego [56] e [57] adota imagens IKONOS tanto multiespectrais quanto pancromática. A partir desses dados o autor desenvolve uma metodologia de classificação baseada em atributos espectrais, contextuais e estruturais para imagens de alta resolução para a área de estudo. Ainda nesse trabalho foram utilizadas regras nebulosas para expressar o conhecimento impreciso do especialista. O autor alcança melhor desempenho através de métodos baseados em segmentos do que por meio de métodos baseados em pixels. Os resultados mostram ainda que a inclusão de conhecimento do especialista através de classificadores baseados em regras nebulosas foi capaz de aumentar significativamente a discriminação de classes que se confundiam no espaço de atributos puramente espectral e estrutural.

Em [58] uma metodologia orientada a objetos foi aplicada para classificar imagens de sensores diferentes, KOMPSAT-1 e SPOT-5. A legenda continha 5 classes de cobertura do solo. O resultado da classificação automática foi comparado com a classificação visual da cena do KOMPSAT-1. Os resultados mostram que a abordagem baseada em objetos não só foi capaz de prover uma diferenciação mais rápida e precisa das classes básicas como foi capaz de, facilmente, se adaptar a outras imagens tomadas em datas diferentes em resoluções espaciais diferentes.

Como objetivo de investigar o método de classificação orientado a objetos na discriminação de diferentes tipos de cobertura de solo em áreas urbanas utilizando imagens de alta resolução, em [59] utilizam-se imagens pancromáticas e multiespectrais IKONOS na cidade de Vienna. A validação baseou-se em dados oficiais, classificados manualmente a partir de ortofotos, por um especialista. Nos experimentos executados obteve-se um erro de 43%, em termos de número de objetos. Por outro lado se analisarmos a área dos objetos classificados corretamente, observamos que essa representa 72% da área total. Essa observação sugere que objetos com áreas menores estão sujeitos a mais erros em termos de classificação do que objetos com áreas maiores.

Imagens SAR (*Synthetic Aperture Radar*) foram analisadas em [60] com o objetivo de desenvolver mapas urbanos de alta resolução. O trabalho foi realizado no eCognition incluindo a segmentação hierárquica em três níveis. Características radiométricas e polarimétricas foram utilizadas em conjunto com características

estruturais como forma e tamanho. Para a validação dos resultados foram utilizados dados oficiais em uma escala compatível com dados SAR. Neste trabalho não foi possível separar algumas classes baseando-se somente nos dados polarimétricos e radiométricos provindos diretamente do SAR, porém quando agregados às características contextuais oriundas do processo de segmentação, como forma ou tamanho dos objetos, as mesmas classes foram distinguidas mais claramente.

Em [38] foi utilizada uma abordagem “mista”, baseada em pixel e em objetos ao mesmo tempo. Inicialmente os pixels da imagem foram classificados utilizando tanto dados espectrais quanto contextuais a partir de um classificador baseado em regras nebulosas. A imagem então foi segmentada e as informações, tais como forma e outros atributos estruturais, foram então adicionadas ao conjunto de regras nebulosas, elaborando uma nova classificação. O desempenho melhorou quanto à taxa de reconhecimento global. Segundo os experimentos somente com a informação de pixels e as regras nebulosas o desempenho é de 10 a 20% maior se comparado a um classificador de máxima verossimilhança. Quando se introduz a análise baseada em objetos o desempenho final aumenta mais ainda, chegando a taxas de reconhecimento de até 86%.

3.3 Classificação Multitemporal

A utilização de imagens multitemporais foi abordada em [61] de forma a representar múltiplas imagens de instantes de tempo distintos como uma série de camadas ou bandas em uma imagem multi-canais. A partir dessa representação, métodos tradicionais baseados em classificadores de máxima probabilidade e dados espectrais foram utilizados normalmente. A estratégia básica no método proposto baseia-se em uma classificação em dois estágios, no primeiro estágio realiza-se as classificações em cada instante de tempo individualmente para que no segundo estágio a decisão final seja tomada através da combinação dos resultados. O artigo propõe três modos de combinação do resultado multitemporal: no primeiro, chamado de adição de probabilidade ou LA (*likelihood addition*) a decisão é tomada para a classe que possui a soma do maior valor nos scores de classificação; o segundo método, denominado de maioria de probabilidade ou LM (*likelihood majority*), atribui a classificação final à classe com maior frequência nas classificações anteriores; o terceiro e último método

utilizado a decisão baseia-se na combinação de regras de Dempster's onde a decisão é tomada onde a medida de aceitabilidade se mostra máxima.

Em [62] o método anteriormente apresentado foi expandido com a utilização de matrizes de coocorrência como atributos espaciais e classificadores baseados em redes neurais artificiais. Nesse trabalho dois tipos de redes neurais foram utilizados, a com treinamento com retropropagação do erro e uma rede neural nebulosa. Dois modelos também foram apresentados: o primeiro baseado em uma rede neural única para a análise das matrizes de coocorrência de todas as imagens avaliadas; o segundo modelo utiliza uma série de redes neurais organizadas em dois estágios, onde cada rede neural no primeiro estágio realiza a classificação da matriz de coocorrência da imagem de cada um dos ano de análise, e então a rede no segundo estágio toma a decisão de classificação baseada nos resultados obtidos das redes do primeiro estágio. Os resultados obtidos mostram um grande incremento de desempenho da tarefa de classificação quando comparado com o método original apresentado em [61].

Em [1] um método de segmentação baseado em conhecimento é proposto. Nesse método é utilizado para a segmentação dados de SIG, conhecimento específico do especialista e dados de série temporais. Nesse trabalho os autores concluem que sistemas baseados em conhecimento utilizados para a tarefa de segmentação oferecem segmentos mais bem delineados e por conseqüência uma classificação mais precisa. Quanto a avaliação multitemporal, o artigo indica que a segmentação usando uma série temporal tende a convergir para uma segmentação ótima conforme o número de imagens da série. Ainda nesse trabalho os autores enfrentam problemas sérios de erro de registro nas imagens de foto aéreas e propõe a utilização de imagens de satélite de alta resolução para resolução desses problemas.

Os dados multiespectrais são tratados em [63] como um conjunto de dados multidimensionais assim como todas as bandas das imagens LANDSAT analisada. O método apresentado extrai atributos do conjunto de dados multidimensionais. A idéia principal é extrair atributos como uma combinação linear de componentes ortogonais de variação dos dados. É demonstrado que com a utilização de dados multitemporais o sistema foi capaz de reconhecer mais facilmente os objetos a serem classificados, ou seja, com menos número de atributos extraídos.

Em [64] e [65] redes neurais artificiais são apresentadas como soluções vantajosas em termos de desempenho global de classificação se comparado com métodos estatísticos.

Em [66] propõe-se um classificador que leva em conta relações estruturais e temporais analisando a interdependência temporal e espacial entre os pixels de uma imagem. O contexto espacial é utilizado para estimar a probabilidade a priori e o conhecimento multitemporal como a probabilidade de transição entre as classes em um determinado intervalo de tempo. O classificador apresentado é testado com imagens bi temporais LANDSAT e mostra um incremento significativo no desempenho da classificação se comparado com métodos tradicionais de máxima verossimilhança.

Em [67] e [68] o conhecimento multitemporal foi empregado por meio de diagramas de transição de estado entre as classes da legenda em um determinado intervalo de tempo. Esses diagramas expressam a possibilidade de transição de uma classe para uma outra classe determinada. Os resultados de ambos os trabalhos mostram que houve um grande incremento no desempenho da classificação em relação à classificação sem informação multitemporal. Porém, em ambos os trabalhos houve dificuldade em modelar as possibilidades de transições entre as classes e foram definidos valores experimentais. Esta abordagem é estendida em [69] e [70] utilizando algoritmos genéticos para estimar os valores dos parâmetros de transição. Isso resultou numa melhora significativa no desempenho de classificação.