

UMA ABORDAGEM MORFOLÓGICA PARA A CARACTERIZAÇÃO DE ESTRELAS E GALÁXIAS

ALCIONE JANDIR CANDÉAS

Algumas das questões de grande interesse para a humanidade são objeto de estudo da Astrofísica extragaláctica, uma das especializações da Astronomia que visa formular idéias básicas com relação à origem e ao destino do Universo. Em sua tarefa de busca para atender a essas questões, muitos projetos de observação do céu implicam a coleção de uma grande quantidade de dados. Entretanto, nem todos os objetos astronômicos atualmente detectáveis pelos instrumentos existentes encontram-se devidamente catalogados. Com o avanço da precisão dos instrumentos de observação e registro nas últimas décadas, muitos objetos estão surgindo nas imagens capturadas sem que se possa proceder com a imediata identificação visual e catalogação.

O problema da discriminação entre estrelas e galáxias é um tema que vem sendo bastante considerado no âmbito das pesquisas que utilizam imagens astronômicas. O sucesso da atividade de discriminação de objetos depende das características selecionadas para identificar as classes de interesse. O trabalho (1) no qual está fundamentado este artigo apresenta uma abordagem original para a caracterização de objetos astronômicos baseada na técnica conhecida como Morfologia Matemática, com um maior destaque nas informações de tamanho e forma em níveis de cinza.

Considerando que o ramo de Processamento Digital e Análise de Imagens (PDAI) tomou um grande impulso na década de 60 devido ao programa espacial da NASA que visava melhorar a qualidade das imagens capturadas das sondas espaciais, este trabalho representa também uma tentativa de reavivar o interesse da comunidade atual de pesquisa em PDAI para aplicações em Astronomia, bem como motivar a multidisciplinaridade.

Introdução

O ramo de Processamento Digital e Análise de Imagens (PDAI) tem sido caracterizado pela crescente adoção de soluções específicas a situações nas quais os métodos que se apresentam apropriados para algumas aplicações podem ser totalmente inadequados para outras. Por outro lado, a Morfologia Matemática (MM) surge como uma teoria geral com uma abordagem unificada para lidar com problemas em muitos campos, como por exemplo: na Medicina, Geologia, Geografia, Sensoriamento Remoto, entre outros. Apesar da diversidade de propósitos, a natureza de todos esses problemas possui uma característica em comum: a necessidade de extrair informações de forma a partir de imagens digitalizadas.

A questão principal no reconhecimento de imagens estelares é se o sucesso de um classificador pode ser aumentado parametrizando a imagem de uma maneira mais eficiente. Com a utilização de ferramentas da MM, buscamos preservar ao máximo as informações de forma dos objetos, onde as características extraídas tendem a ser adequadas para a caracterização de tais objetos astronômicos, ao mesmo tempo em que também se propõem efetivas para uma possível futura classificação automática baseada, por exemplo, em Redes Neurais artificiais.

Considerações em Astronomia e Astrofísica

Uma constatação interessante, feita por muitos historiadores, é que as primeiras ciências a se desenvolverem foram justamente as que tratam do que está mais distante do homem, como por exemplo, a Astronomia. Se se pensar em termos de observação, certamente a Astronomia é a mais antiga das ciências, e surgiu com a associação feita a partir da ocorrência de certos fenômenos naturais e o movimento peculiar dos astros, ou corpos celestes. No início, tais observações limitavam-se ao propósito de medição do tempo, atividades agrícolas, cultos à natureza, etc..

Com o passar do tempo, a Astronomia foi evoluindo junto e com o auxílio da Matemática. Outras ciências foram surgindo: exatas, humanas, tecnológicas, sociais, etc.. Atualmente, olhar para os astros não se restringe apenas aos astrônomos. Poetas se inspiram com a beleza e harmonia do Universo; cidadãos comuns admiram o infinito...; o homem já consegue prever fenômenos; as funções do cotidiano das vidas humanas giram em torno da noção de dia/noite; tira-se proveito

do conhecimento a respeito da energia solar, da força magnética da lua sobre as marés, e tecnologias e resultados alcançados ao longo das pesquisas espaciais têm sido amplamente utilizados para aperfeiçoar e curar.

Com o desenvolvimento das pesquisas astronômicas, ramificações foram surgindo de acordo com as áreas específicas de interesse. A estrutura e dinâmica galáctica, os efeitos ambientais na organização e evolução galáctica, a diversidade de tipos morfológicos de galáxias, além da distribuição de matéria clara e escura em larga escala por todo o universo, são alguns dos problemas relevantes discutidos na Astrofísica, que compreende uma das ramificações da Astronomia atual.

O conhecimento da estrutura dos Aglomerados - concentrações mais densas de galáxias normalmente encontradas no Universo - e do seu grau de agrupamento também pode fornecer informações essenciais para o entendimento do arranjo entre as galáxias e dos processos dinâmicos dentro dessas estruturas, além da distribuição da matéria e informações a respeito da existência de galáxias de tipos morfológicos distintos. Tais problemas cosmológicos são normalmente estudados a partir de estatísticas para tratar amostras de uma grande quantidade de dados de estrelas e galáxias que são coletados sobre extensas regiões do céu.

A maioria das informações que sabemos a respeito do Universo deve-se, principalmente, à luminosidade proveniente do espaço. Inúmeras foram as ferramentas utilizadas pelo homem ao longo de sua evolução, para registrar os fenômenos observados com o movimento dos astros no céu, partindo de instrumentos rudimentares de pedra e metal para gravar ou esculpir o que a visão tivesse a oportunidade de presenciar. A forma de registro foi evoluindo, porém a abordagem dos fenômenos astronômicos do dia-a-dia era essencialmente descritiva, uma abordagem Cinemática, sem se ocupar com a causa dos fenômenos. Antes da luneta de Galileu era preciso valer-se de argumentos complicados para proporcionar compreensão de alguns fenômenos, como por exemplo, na observação de fases em Vênus. Com Newton, surgiu uma ferramenta poderosa e com mais significado físico: a Dinâmica, que aliada à observação, veio explicar os movimentos dos astros.

A Astronomia moderna deve a maior parte de suas realizações ao telescópio e à câmera - as principais ferramentas dos astrônomos. Em seus numerosos estudos e aplicações práticas, muitos cientistas também estão familiarizados com a utilização de catálogos estelares,

tais como o *Yale Bright Star Catalog* e o *SAO XV Catalog*. Alguns dos catálogos mais completos foram compilados visualmente a partir de amostras fotográficas (2).

Com o advento de novas ferramentas de medição espectroscópica terrestres (ou seja, manipuladas na superfície da Terra), aliadas ao lançamento de observatórios espaciais como o *Hubble Space Telescope*, catálogos de estrelas e galáxias mais profundos e mais precisos são exigidos, pois mais objetos têm surgido nas imagens, devido não só ao aumento no alcance, mas também na resolução dos dados.

A tendência atual é se utilizar ferramentas de detecção de alta precisão e máquinas de *scanning* rápidas para compilar catálogos de galáxias sobre áreas específicas do céu.

Imagens Astronômicas

O desenvolvimento de métodos e ferramentas de aquisição de imagens, tais como as câmeras CCD (*Charge Coupled Device*) e *scanners* especiais para as placas fotográficas acopladas aos telescópios de grande porte, vem proporcionando o aumento dos dados disponíveis na forma digital, impulsionando a relevância da interação com a computação para servir a propósitos astronômicos no sentido de poder aperfeiçoar os resultados dos programas observacionais. No entanto o correspondente acréscimo dos dados nas imagens devido à rapidez e precisão de tais instrumentos prescinde de novas técnicas para detecção e classificação automática dos objetos.

A manipulação de imagens astronômicas digitalizadas envolve vários aspectos que devem ser levados em consideração, tais como: compressão dos dados sem perda de informação, utilização de processamento paralelo para acelerar o tratamento das imagens com diferentes resoluções, conhecimento dos tipos de ruído conforme o processo de aquisição das imagens, além de uma gama de outros problemas de análise de imagens que são familiares aos pesquisadores nessa área.

Uma notável contribuição para os *surveys* (amostras) de estrelas e galáxias é o conjunto de imagens digitalizadas provenientes dos 936 pares de placas fotográficas do *National Geographic Palomar Observatory Sky Survey* (POSS). Tais imagens foram extraídas utilizando o *Automated Plate Scanner* (instalado no Monte Palomar, EUA) do Projeto e Catálogo APS (3), da Universidade de Minnesota.

Uma única placa do POSS chega a produzir aproximadamente 250.000 objetos detectados.

Uma outra contribuição às amostras de imagens astronômicas é o *Digital Sky Survey* (DSS), que consiste em uma coleção de amostras de imagens abrangendo toda a esfera celeste e é amplamente utilizada em trabalhos de análise e classificação pela comunidade astronômica nacional e internacional. Trata-se de um conjunto de 41 *CD-ROMs* produzidos pelo *Catalogs and Surveys Branch* do *Space Telescope Science Institute* (STScI), contendo um total de 645 imagens, uma base de calibração astrométrica e um *software* para acesso aos dados.

As amostras do hemisfério Norte celeste do DSS são baseadas nos dados fotográficos do POSS, enquanto as do hemisfério Sul celeste são provenientes do *SERC Southern Sky Survey*, produzidas no *Anglo-Australian Observatory* (os direitos autorais do material original das placas fotográficas pertencem ao *Royal Observatory Edinburgh* e ao *Anglo-Australian Observatory*, na Austrália, a partir de um telescópio UK Schmidt 1.2 m - uma câmera de propósito especial com um amplo campo de visão). As placas fotográficas foram processadas na forma digital comprimida (técnica *H-transform wavelet*) e foram inicialmente geradas para a construção do Catálogo GSC (*Guide Star Catalog*).

Todas as amostras estão disponíveis como imagens FITS (*Flexible Image Transport System*) - formato de dados padrão para intercâmbio e armazenamento, empregado internacionalmente pelos pesquisadores na área de Astronomia.

O Problema da Discriminação Estrela/Galáxia

Devido ao aumento na precisão e na resolução das ferramentas de registro e observação nas últimas décadas, não há possibilidade de catalogar apropriadamente e em tempo hábil todos os objetos atualmente detectáveis, pois, como dito anteriormente, a quantidade de objetos vem aumentando progressivamente com a evolução de tais ferramentas. A catalogação automática de objetos astronômicos requer uma solução em PDAL. Isso deve-se principalmente ao processo de catalogação requerer, normalmente, a identificação visual, tornando-se, com todos esses novos aspectos, uma tarefa cansativa e desgastante. Sabendo-se que uma simples placa fotográfica pode produzir cerca de milhares de objetos detectados, torna-se necessário o desenvolvimento de procedimentos automáticos para processamento rápido das imagens, reconhecimento e classificação dos objetos.

Uma das questões mais importantes em se tratando do trabalho de catalogação automática é o problema da discriminação entre estrelas e galáxias, que é atualmente bastante considerado na identificação automática de objetos astronômicos. Esse tema tem sido abordado por diversos grupos de pesquisa e profissionais (4), considerado como um problema bastante complexo em relação a muitos outros existentes, por vários motivos, dentre os quais podem ser destacados: quantidade excessiva de objetos em uma imagem, presença de ruídos frequentes, grande quantidade de objetos pequenos indefinidos, diferentes caracterizações de objetos de mesma classe porém de tamanhos diferentes, entre outros.

Uma contribuição interessante nesse sentido é o APS *Project and Catalog*, que tem gerado um catálogo astronômico disponível na Internet (<http://aps.umn.edu/>) para poder ser manipulado por outros grupos, com milhões de estrelas e galáxias classificadas por técnicas de reconhecimento de padrões para realizar a distinção estrela/galáxia em grandes *surveys* digitais do céu. A implementação é baseada em calibração fotométrica e normalização de conjuntos de parâmetros considerados críticos na imagem, utilizando classificadores baseados em redes neurais para tentar fazer, de forma automática, a discriminação esperada.

No entanto, todos os catálogos possuem limitações devido aos critérios adotados. Para o pesquisador, necessário se faz verificar a qualidade do catálogo utilizado com relação à completude e corretude a fim de produzir resultados de análises confiáveis. Essas limitações dependem de vários fatores, não somente da precisão do instrumento de aquisição correspondente à imagem escolhida, mas também (e principalmente) das características selecionadas para serem extraídas de cada objeto - que devem representar um padrão para posterior reconhecimento -, além da confiabilidade do classificador empregado.

Uma vez garantida a qualidade da imagem, ferramentas de análise adequadas são necessárias para que se possa extrair informações relevantes dos dados. O escopo do problema de quantos e quais parâmetros são necessários para uma caracterização precisa de um simples objeto astronômico é um tópico bastante importante a ser considerado. Esse assunto vem sendo muito discutido ultimamente pela comunidade de pesquisa em Astronomia.

Discriminação entre Estrelas e Galáxias

A maioria dos trabalhos encontrados na literatura relacionados à identificação de objetos astronômicos em imagens do céu digitalizadas são baseados em características extraídas a partir de atributos fotométricos. Isso significa que tais técnicas empregam as intensidades dos *pixels*, que estão diretamente relacionadas ao brilho aparente do objeto. Em tais trabalhos uma grande variedade de métodos tem sido empregada a fim de alcançar uma parametrização efetiva das imagens, combinando os atributos fotométricos com alguma medida do tamanho do objeto, como por exemplo o diâmetro, aliado ao cálculo de simples momentos e gradientes.

Entretanto imagina-se inicialmente que, como a maioria dos catálogos existentes foram baseados na identificação visual, quanto mais assemelhar-se a esse processo, mais confiável tende a ser qualquer proposta de classificação automática a ser realizada.

Esta pesquisa propõe uma abordagem morfológica para a caracterização dos objetos astronômicos, que se torna intuitivamente mais aproximada à maneira pela qual o ser humano realiza a análise visual das imagens astronômicas.

Foram utilizadas, para atingir o objetivo proposto, ferramentas da Morfologia Matemática para extrair informações de forma em níveis-de-cinza, que consideramos apropriadas para a discriminação entre estrelas e galáxias em imagens de *surveys* de galáxias em larga escala, como as do DSS e do POSS. Esses tipos de imagens contêm, tipicamente, milhares de objetos astronômicos de tamanho e brilho variados, a maioria estrelas e galáxias, conforme pode ser visto na Figura 1 (esta amostra de imagem foi invertida com propósitos de contraste, para uma melhor visualização).

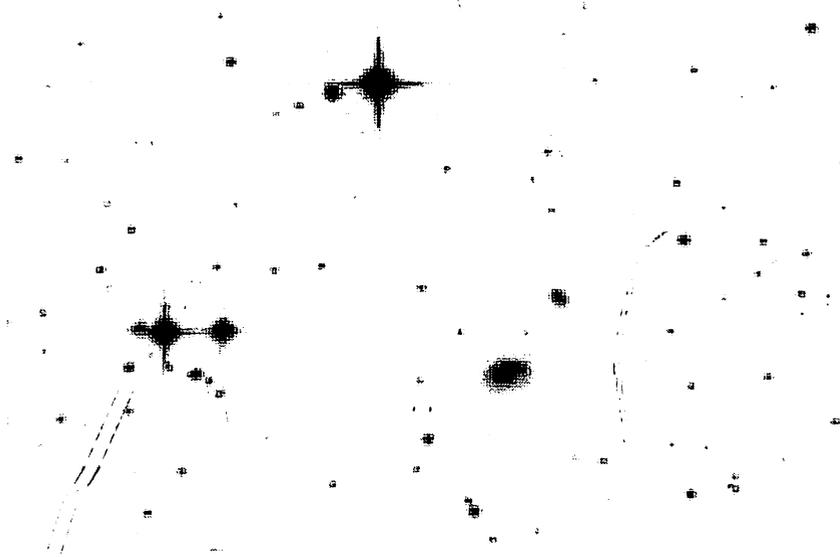


Figura 1 - Amostra de imagem de uma concentração de galáxias em larga escala - extraída do DSS

Tanto a inspeção visual de *surveys* astronômicos, quanto a automática, têm provado serem tarefas de grande dificuldade. Um observador humano pode perceber, nesse tipo de imagem, que as estrelas são objetos bem definidos, enquanto que as galáxias se apresentam como manchas um pouco "borradas". Além do mais, estrelas "grandes" apresentam, nas amostras do DSS, um efeito de cruz devido às propriedades óticas do instrumento de aquisição.

Um outro aspecto que se observa é que estrelas e galáxias de mesmo tamanho aparente possuem comportamentos distintos no que se refere à distribuição de *pixels* significativos (ou seja, de níveis de cinza mais representativos em um objeto). As estrelas apresentam-se mais compactas, com uma alta densidade de *pixels* por todo o objeto, enquanto as galáxias possuem um núcleo mais denso, porém uma extensa área em torno na qual a densidade de *pixels* vai diluindo com a distância do centro. Neste trabalho estamos propondo utilizar tais características específicas de forma para discriminar entre essas duas classes de objetos, por acreditarmos representarem o principal aspecto considerado ao longo do processo de identificação visual.

Morfologia Matemática em Níveis de Cinza

Muitos fenômenos comuns podem ser modelados como funções lineares. Além disso, métodos lineares freqüentemente transformam um problema em outro mais simples no domínio da freqüência. Contudo, isso representa uma restrição matemática que pressupõe o tipo de processamento que se pode realizar, podendo levar a efeitos colaterais indesejáveis. Mesmo com restrições próprias, a não-linearidade pode ser mais apropriada para uma aplicação em particular. Por exemplo, se a informação de forma é fundamental em uma aplicação, deve-se procurar aplicar restrições baseadas na forma.

Em PDAI o interesse reside em tarefas complexas, tais como descobrir as distribuições de tamanho da imagem ou classificar objetos. A principal vantagem dos métodos não-lineares é a capacidade de preservar, de forma seletiva, informações estruturais enquanto realizam alguma tarefa na imagem.

A Morfologia Matemática (MM) refere-se a uma ramificação do processamento não-linear e análise de imagens que está centrada nas estruturas geométricas contidas em uma imagem. Seus fundamentos teóricos residem na Teoria dos Conjuntos e na Teoria de Reticulados. A MM nasceu na década de 60, na França, com o trabalho de Georges Matheron e Jean Serra, sendo criada na tentativa de reparar o desenvolvimento empírico dos filtros não-lineares que mais se aproximam do propósito do processamento de sinais (caracteristicamente linear, diferentemente do processamento de imagens). Desde então, o processamento morfológico de imagens tem evoluído bastante por todo o mundo, tanto nos estudos teóricos como nas aplicações das técnicas desenvolvidas.

A teoria original restringia-se à aplicação em imagens binárias a partir dos resultados inicialmente formalizados para os reticulados dos subconjuntos, mas com os avanços teóricos, os mesmos resultados foram estendidos para o reticulado das funções, tornando-se evidente que as imagens em níveis de cinza poderiam ser tratadas algebricamente como no caso binário, estendendo para o domínio dos reticulados completos quaisquer. A partir de então, foi desenvolvida uma fundamentação teórica em torno da Morfologia em níveis de cinza, estendendo os princípios da MM de duas para três dimensões, baseando-se na noção intuitiva de imagens em tons contínuos sendo visualizadas como superfícies tridimensionais, cujas topologias podem ser modificadas "investigando-as", ou seja, percorrendo-as com estruturas geométricas.

Quando se trabalha com imagens em níveis de cinza, é comum visualizar uma função de uma imagem em perspectiva com um terceiro eixo representando o brilho, ou valor de cinza, onde as regiões com grande variação de brilho aparecem com uma série de elevações ou depressões, e regiões suaves ou de planaltos onde os níveis de brilho variem pouco ou sejam constantes. Sendo assim, uma imagem digital é uma imagem $f(x,y)$ que foi discretizada tanto em coordenadas espaciais como em brilho. Se o fundo da imagem tende a ser escuro, os objetos da imagem devem ser visualizados no modelo topográfico como picos ou elevações, e como vales ou depressões para o caso do fundo da imagem ser mais claro.

Os princípios da MM podem ser estendidos a conjuntos no espaço Euclidiano n -dimensional. Os operadores da MM em nível de cinza utilizados neste trabalho lidam com funções $f: E \rightarrow [0, k] \in \mathbb{Z}$, onde E corresponde ao *grid* planar digital usual, ou seja, uma matriz representando alguma porção do plano cartesiano cujos pontos correspondem às posições dos *pixels* na imagem. f é uma função em nível de cinza situada nos pontos do espaço Euclidiano bidimensional, denotando os valores em níveis de cinza em cada ponto, dando uma noção tridimensional onde, como introduzido anteriormente, se o fundo da imagem for em um tom mais escuro que os objetos, a imagem assemelhar-se-á a uma representação em alto-relevo. Caso o oposto ocorra, ou seja, os objetos forem em tons mais escuros que o fundo da imagem, a semelhança dar-se-á uma representação em baixo-relevo. Isso corresponde ao conceito usual de imagens em níveis de cinza.

A idéia geométrica básica é percorrer a imagem investigando-a com uma função definida em um pequeno subconjunto do domínio de definição da imagem - o elemento estruturante -, a fim de extrair informações de tamanho e forma através de transformações, verificando a maneira na qual ele se "encaixa" (está ou não contido) na imagem.

Operadores em Níveis de Cinza

Na Morfologia em níveis de cinza, as funções mínimo e máximo também possuem importantes papéis em contrapartida à interseção e união na Morfologia Binária. Os operadores duais básicos da MM em níveis de cinza são a dilatação e a erosão, definidos, respectivamente, por (5):

$$\delta_g(f)(x,y) = \max \left\{ f(x',y') + g(x-x',y-y') \right\},$$

$$\forall (x',y') \in (B' + (x,y)) \cap E$$

$$\varepsilon_g(f)(x,y) = \min \left\{ f(x',y') - g(x-x',y-y') \right\},$$

$$\forall (x',y') \in (B + (x,y)) \cap E$$

onde f corresponde à imagem e g e B denotam, respectivamente, os valores do elemento estruturante e do domínio de definição ($B \subseteq E$). B' é necessário para tornar a dilatação dual à erosão. Os operadores de adição e subtração utilizados acima não são os utilizados normalmente nas operações comuns. Tratam-se dos operadores de Heijman, que consistem de versões sutilmente modificadas, necessárias para atender a restrições algébricas (6), garantindo que os resultados permaneçam dentro de um intervalo pré-definido.

A composição dos operadores de erosão e dilatação resulta, ainda, nos operadores morfológicos básicos de abertura e fechamento, que são definidos, respectivamente, por:

$$\gamma_g = \delta_g \varepsilon_g$$

$$\phi_g = \varepsilon_g \delta_g$$

Aberturas e fechamentos são filtros morfológicos (7) que possuem excelentes propriedades de remoção de ruídos. A partir do modelo topográfico tridimensional pode-se identificar que aberturas e fechamentos em níveis de cinza, selecionando-se apropriadamente o tamanho e a forma do elemento estruturante, possuem a propriedade de remover detalhes da imagem - a abertura vai minimizar "picos" em uma configuração de alto-relevo e o fechamento taparia "buracos" em uma configuração de baixo-relevo. O operador de fechamento é o mais freqüentemente utilizado neste trabalho, uma vez que ele age diretamente sobre o fundo em níveis de cinza com valores mais altos considerando que as imagens foram invertidas para atingir o efeito desejado na implementação.

Ferramentas Morfológicas

Muitas transformações nas imagens que se quer trabalhar podem ser derivadas a partir dos operadores morfológicos básicos para alcançar um resultado específico. A seguir serão apresentadas as ferramentas morfológicas que foram utilizadas para a implementação desta abordagem.

Top-Hat: Uma ferramenta bastante apropriada para realce pode ser obtida a partir de aberturas ou fechamentos, conhecida como transformação *top-hat*. Ela possui a propriedade de realçar "picos" (elevações) ou "vales" (depressões) aplicando, respectivamente, o operador de abertura para o caso do fundo da imagem ser escuro, ou fechamento no caso inverso, ou seja, quando os objetos forem mais escuros que o fundo da imagem. Os operadores da transformação *top-hat* para os casos de abertura e fechamento são definidos, respectivamente, por:

$$\begin{aligned}\tau_{\text{og}}(f) &= f - \phi_g(f) \\ \tau_{\text{cg}}(f) &= \phi_g(f) - f\end{aligned}$$

Após suavizar a imagem com a operação de abertura (respectivo fechamento), a subtração pela imagem original fornece o resultado de realce. Devido à simplificação do operador de fechamento, a *top-hat* possui a propriedade de realçar "vales", que são regiões mais escuras na imagem, servindo apropriadamente como um bom removedor de fundo da imagem.

A escolha do tamanho do elemento estruturante é muito importante. Se ele for muito pequeno, a remoção do fundo é afetada pela presença de ruído aleatório. Se, ao contrário, for muito grande, não será possível reproduzir as variações de escala dos objetos fracos. Sendo assim, é necessário que se encontre um compromisso na seleção desses valores.

Watershed: A transformação *watershed* (ou "bacia-d'água") é uma poderosa ferramenta da MM para segmentação (8). Essa metodologia possui uma série de variações, porém a abordagem mais intuitiva é pensar em uma imagem como sendo um modelo topográfico onde os vales possuem mínimos regionais, e supor que água está sendo colocada nos vales (ou bacias) de maneira que vão sendo enchidos e a

água vai se espalhando a uma mesma velocidade a partir de cada mínimo regional, iniciando do mais baixo e então a partir dos demais cada vez que o nível da água atingir suas altitudes. Diques são construídos nos lugares onde as águas provenientes de diferentes mínimos se encontram, separando, assim, as "bacias-d'água". Os diques que ficam acima da superfície da água constituem as linhas divisórias de *watersheds*, que consistem em contornos fechados que envolvem cada um dos mínimos regionais (onde estariam os prováveis objetos) e correspondem às cristas do relevo, representando uma boa segmentação.

Última Erosão: O conceito de última erosão está relacionado diretamente a imagens binárias e é proveniente de sucessivas erosões de uma imagem por um único elemento estruturante. Tomando a imagem como sendo constituída de subconjuntos que se sobrepõem, a cada passo os componentes conexos de *pixels* podem ser reduzidos, separados ou desaparecerem. Os resíduos resultantes de cada componente constituem a última erosão da imagem e são freqüentemente utilizados como conjuntos de marcadores para processamentos posteriores, tais como rotulação, reconstrução ou identificação de partículas sobrepostas (9).

Granulometria: Ao se utilizar uma família adequada de fechamentos (resp. aberturas) obtém-se uma poderosa ferramenta de descrição de tamanho e forma para análise de imagens, conhecida como granulometria, método desenvolvido por Matheron para análise de imagens "granulares" (10). Ao se pensar em uma imagem como uma coleção de "grãos", para que um grão em particular passe ou não através de uma peneira, irá depender de seu tamanho e forma relacionados com a forma da malha da peneira. Ao se aumentar, sucessivamente, o tamanho dessa malha enquanto a sua forma básica é mantida, mais conteúdo da imagem vai passando em cada passo. A tendência é que ao final não mais permaneça grão algum. Assim, pode-se dizer que o método baseia-se na "peneiração" da imagem acompanhada de medições sucessivas da quantidade de resíduo que permanece na peneira.

Na prática, a granulometria consiste de uma seqüência de fechamentos (resp. aberturas) por um conjunto de elementos estruturantes apropriadamente selecionados, cada vez maiores, obedecendo a uma ordem crescente e pertencendo à mesma família, isto

é, mantendo a forma básica, que é dita como geradora da granulometria. Ao medir-se o volume sob a imagem após cada fechamento (resp. abertura), pode-se construir uma curva de distribuição de tamanho:

$$\Phi(\lambda) = \frac{V(\lambda) - V(0)}{V(\Lambda) - V(0)}, \lambda \geq 0$$

onde λ é a parametrização da família de elementos estruturantes utilizados nos fechamentos, $V(\lambda)$ é o volume do fundo da imagem (i. e., a integral positiva) a cada iteração e Λ é o parâmetro associado ao maior elemento estruturante, selecionado por ser aquele largo o suficiente para eliminar, ou fechar completamente o objeto de interesse. Note que Λ é, por si só, um atributo de análise bastante útil.

Para alcançar um valor preciso do resíduo deixado, é necessário que se subtraia o volume inicial $V(0)$ do último volume encontrado. A idéia subjacente ao resíduo variando a cada iteração tem relação direta com o tamanho e a forma do objeto. Ao dividir pelo resíduo a ser atingido, a função acima torna-se monotonicamente crescente dentro do intervalo $[0,1]$, de maneira que pode ser vista como uma distribuição de probabilidade cumulativa. Sua função densidade de probabilidade associada é conhecida como o espectro de padrões da imagem relativo à granulometria e deve, idealmente, tornar-se nula para $\lambda > \Lambda$, o que representaria não haver mais resíduo.

O espectro de padrões é uma poderosa ferramenta para caracterização do tamanho e da forma em análise de imagens (11). De uma maneira reminescente ao espectro de Fourier, mostra a decomposição de um dado objeto em termos de uma forma fundamental em escala conforme os valores crescentes do parâmetro λ .

Alguns atributos de análise bastante úteis que podem ser definidos baseados no espectro de padrões são o raio médio e a rugosidade média. A rugosidade média é a entropia da teoria da informação do espectro de padrões, a função de incerteza que quantifica a complexidade de tamanho-forma do objeto, ou seja, como a forma do objeto difere da forma esperada.

Procedimentos da Abordagem Morfológica

As imagens FITS utilizadas neste trabalho foram adquiridas a partir do *survey* em grande-escala do DSS. Selecionamos imagens representativas que possuíssem uma grande quantidade de galáxias:

- Group 1701: α [02 08 32.1] δ [-55 39 6.5]
- ABELL 3667: α [20 10 50.3] δ [-56 40 23]
- ABELL 3698: α [20 35 08.1] δ [-25 15 34]
- ABELL 3775: α [21 31 24.8] δ [-43 18 38]
- OO140: α [01 04 34.3] δ [-23 50 08]

onde α e δ representam as coordenadas celestes estimadas para o ano 2000, sendo α a *ascensão reta* em horas, minutos e segundos, e δ a *declinação* em graus, minutos e segundos. Cada uma dessas imagens consiste em aproximadamente 500X500 *pixels* (correspondendo, no DSS, a 15'x15' arcsec no céu).

O processamento de uma imagem pode ser abordado em diferentes etapas. A seguir serão descritas as etapas consideradas neste trabalho.

Pré-processamento: A etapa de pré-processamento concentrou-se na extração do fundo de céu, tentando-se corrigir uma certa quantidade de imperfeições decorrentes da aquisição da imagem.

Inverter a imagem é uma boa prática ao se lidar com imagens astronômicas, cujo propósito é proporcionar um maior realce do contraste nos objetos, visando facilitar a identificação pelos observadores visuais (naturais). Neste trabalho houve também uma outra razão pela opção em utilizar a imagem invertida: embora muitos dos operadores da MM sejam duais com relação ao complemento em nível de cinza, a inversão é necessária para obtenção do modelo topográfico ideal da imagem (onde os objetos de interesse tornam-se "vales") requerido pela abordagem da *watershed* na etapa de segmentação.

No tratamento de imagens astronômicas de grande escala, frequentemente faz-se necessário realçar a imagem, diminuindo, assim, efeitos provenientes das condições ambientais (luminosidade, densidade atmosférica, entre outras) da superfície no momento da aquisição da imagem, ou até das características específicas dos instrumentos (por

exemplo, efeitos óticos como halos em torno de “grandes” estrelas) que podem vir a afetar o processo de segmentação, corrompendo objetos no interior dessas regiões, sugerindo tratar-se de um objeto extenso, além de modificar a aparente forma de cruz das estrelas que se deseja evidenciar em uma imagem típica do DSS. Essa operação de remoção do fundo da imagem e realce dos objetos pôde ser bem realizada a partir de uma transformação *top-hat* da imagem invertida, utilizando como elemento estruturante um disco Euclidiano planar (uma aproximação digital do círculo) de raio $r=8$. Esse valor fez-se necessário para garantir a remoção do halo da maior estrela nas amostras selecionadas para o nosso trabalho.

A composição de uma abertura seguida de um fechamento é utilizada para suavizar o fundo da imagem, prevenindo contra uma super-segmentação, removendo o ruído de fundo (“picos”) que é realçado pela aplicação da transformação *top-hat*. Foi utilizado o elemento estruturante planar, que consiste em uma função de valor zero possuindo apenas seu domínio de definição, ou seja, um conjunto. Neste trabalho foi utilizado o quadrado 3X3.

Segmentação: Uma boa prática para se otimizar o resultado da aplicação do método de segmentação *watershed* consiste em regularizar a imagem modificando sua homotopia (12). É necessário que se obtenha marcadores apropriados, ou conjuntos conexos de *pixels*, para identificar cada objeto individual que se deseja extrair. Isso envolve conhecimento *a priori* das imagens em estudo. Um marcador de fundo também é requerido.

Neste trabalho, inicialmente tentou-se encontrar um marcador que preservasse a maioria dos objetos astronômicos mais fracos. Uma solução imediata consistia em marcar manualmente os objetos, mas uma extração automática dos marcadores era preferível, dado a grande quantidade de objetos em cada imagem.

Ao se utilizar um limiar, obtém-se marcadores que são razoavelmente insensíveis a ruído residual de fundo na imagem realçada. Em nosso caso, foi estimado um valor de limiar (210), destacando os objetos mais significativos e preservando objetos fracos. Todavia tal escolha ainda levou a situações onde um único marcador foi encontrado para objetos que se sobrepunham.

Ao aplicar o operador da última erosão após o limiar, levou a marcadores mais satisfatórios. Ao aplicar a *watershed* diretamente à imagem limiarizada, levou ao fornecimento de linhas divisórias das regiões, que puderam ser utilizadas como um marcador do fundo. Uma

vez extraídos os marcadores, um procedimento morfológico permitiu impor esse conjunto de marcadores para criar mínimos regionais em suas localizações.

Na interpretação da *watershed* como enchimento de regiões em depressões, a saída é uma tesselação da imagem, onde cada contorno fechado, rotulado por um nível de cinza diferente, define uma região que provavelmente contém um objeto. Com a aplicação da modificação de homotopia a partir da imagem resultante da *top-hat*, foi possível tratar as regiões vazias na segmentação.

Na Figura 2 pode-se visualizar uma das imagens trabalhadas e o resultado das etapas descritas anteriormente.

Extração de Características: Em aplicações como essas, um passo fundamental para proceder com a extração de características após a segmentação, consiste em isolar apropriadamente, cada região de interesse centralizada em uma imagem de vizinhança local, para que cada objeto possa ser tratado separadamente para cálculo dos atributos de forma. Na nossa abordagem morfológica implica que o espectro de padrões e medidas associadas de raio médio e rugosidade referentes ao objeto sejam computados. Para que os efeitos de borda fossem minimizados no cálculo, foi necessário uniformizar a região com o valor médio da imagem original inteira, como pode ser visto na Figura 3.

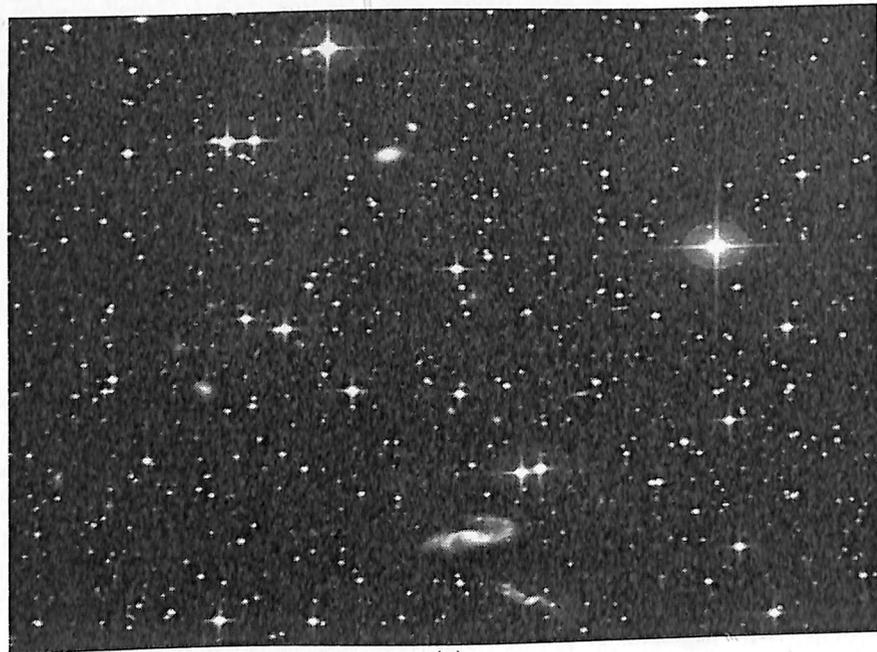
Para a descrição dos objetos selecionados, foi explorado o conceito de granulometria. Assim, partiu-se para a escolha dos elementos estruturantes que melhor definissem a estelaridade de um objeto. Foram consideradas, inicialmente, três famílias diferentes de elementos estruturantes tridimensionais básicos, associados aos discos digitais não-planares de raio crescente baseados nas métricas do quarteirão, euclidiana e do tabuleiro, que consistem, respectivamente, na aproximação digital de uma pirâmide de base quadrada, uma semi-esfera e um cubo.

Resultados Experimentais

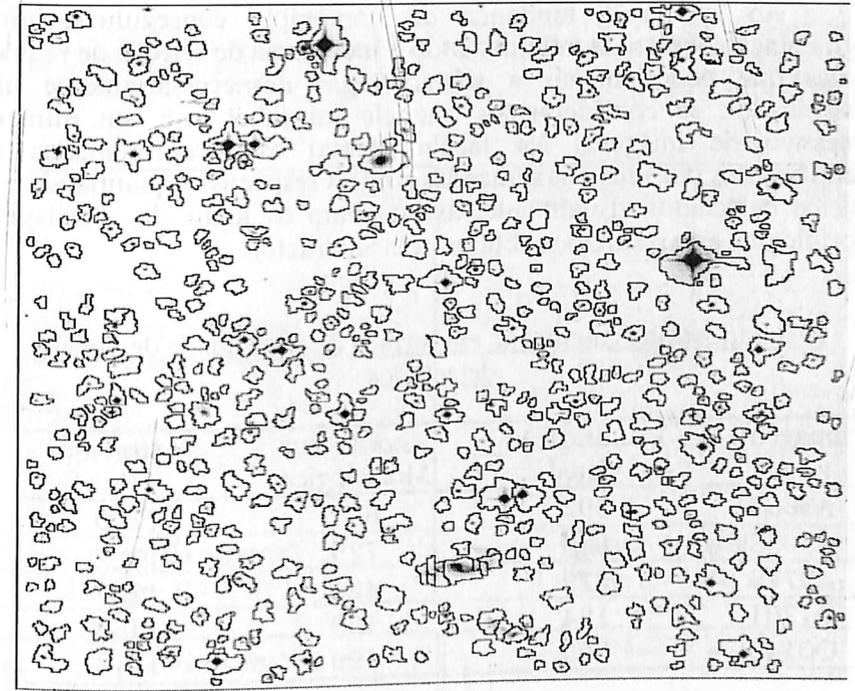
Comparamos os resultados de nossa abordagem com outras soluções já estabelecidas, mais especificamente com o trabalho desenvolvido no Instituto Astronômico e Geofísico da USP (IAG) para o tratamento de imagens digitais contendo aglomerado de galáxias objetivando determinar os parâmetros fotométricos e elípticos das galáxias presentes nas imagens.

Os resultados da segmentação são comparados com os do SExtractor, um *software* desenvolvido para lidar especificamente com imagens astronômicas, mas que não foi capaz de separar convenientemente objetos sobrepostos. Outro problema típico de segmentação inclui a fragmentação de galáxias Sc extensas devido à característica própria da não-homogeneidade de seus braços em espiral (uma galáxia compõe-se de um grande número de objetos).

A imagem de entrada no SExtractor foi produzida utilizando o IRAF (*Image Reduction Analysis Facility*), *software* gráfico desenvolvido pelo grupo *Iraf Project – National Optical Astronomy*



(a)



(b)

Figura 3 - (a) imagem original ABELL3698; (b) visualização da segmentação resultante da seqüência de operações da MM na imagem original

Observatories (NOAO). Os procedimentos de pré-processamento e classificação visual realizados pela equipe do IAG comumente utilizam o IRAF, que é amplamente utilizado pela comunidade internacional de profissionais de Astronomia e Astrofísica.

A abordagem morfológica foi implementada no sistema de processamento e visualização de imagens conhecido como Khoros, em plataforma Unix. O Khoros é um ambiente aberto de PDAI desenvolvido na Universidade do Novo México e mostrou ser uma ferramenta bastante conveniente para desenvolvimento de algoritmos, integração de sistemas e rápida prototipação de aplicações devido ao seu ambiente de programação visual - o Cantata. Um conjunto de programas desenvolvidos no *Cantata* pode ser agrupado e organizado como uma "caixa-de-ferramentas" (*toolbox*) que pode ser facilmente integrado ao sistema de acordo com a necessidade do usuário. Neste trabalho foram utilizados os operadores da *toolbox* MMach (13).

Ao aplicar a mudança de homotopia conseguiu-se uma segmentação otimizada, minimizando a incidência de seleção de regiões vazias que poderiam vir a sobrecarregar desnecessariamente um classificador, se considerarmos que ele irá lidar com um número excessivo de imagens. Na tabela abaixo estão apresentadas as identificações das imagens utilizadas com as respectivas quantidades de objetos detectados visualmente, assim como os totais da abordagem morfológica e os resultados obtidos pelo SExtractor.

Resultados da Segmentação em termos de quantidade de objetos detectados

Imagem	Identificação Visual	Abordagem Morfológica	Sextractor
A3667	592	651	1700
A3698	688	725	1722
A3775	379	421	1257
G1701	184	159	1457
O0140	209	199	1413

Os melhores resultados da extração de características foram obtidos com a métrica do quarteirão. Na Figura 4 podem ser vistas 4 objetos com diferentes tamanhos aparentes e seus respectivos espectros de padrões e atributos de análise associados.

O espectro de padrões diferencia-se nitidamente nos objetos "maiores", identificados pelos maiores raios. A estrela "grande" produz uma curva bimodal devido ao efeito de cruz, enquanto que a galáxia está associada com uma curva unimodal devido à granulometria uniforme. O primeiro "pico" no espectro de padrões da estrela "grande" corresponde ao processo de remoção das extremidades da cruz, caracterizando bem a complexidade relacionada à forma geométrica investigada (métrica do quarteirão). Isso também reflete no valor da rugosidade média (θ), que é maior para a estrela. Para os objetos "menores", o espectro de padrões da estrela ascende de maneira mais rápida, como resultado da transição abrupta do nível de cinza do fundo para o objeto.

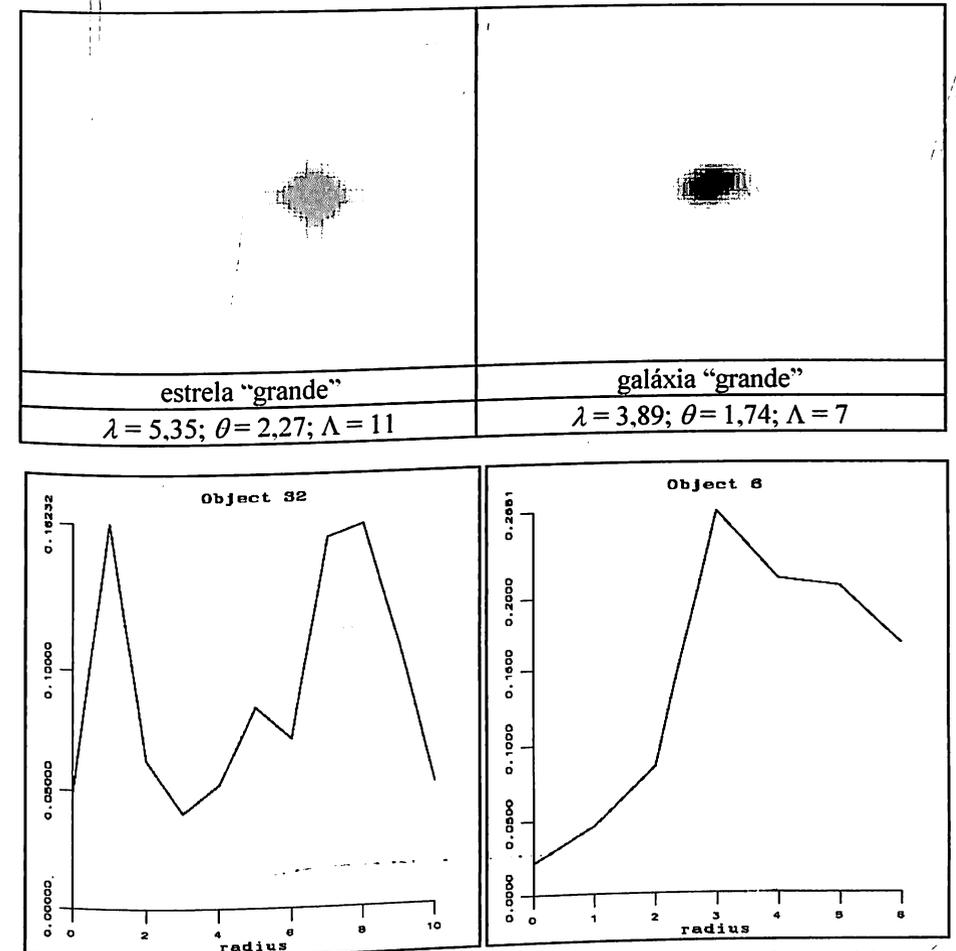


Figura 4 - Espectro de Padrões e atributos associados de 4 objetos utilizando a métrica do quarteirão

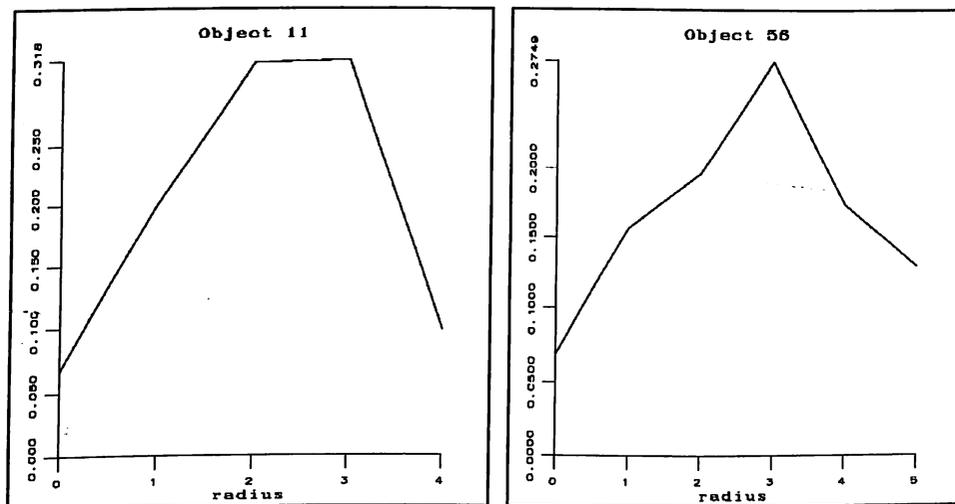
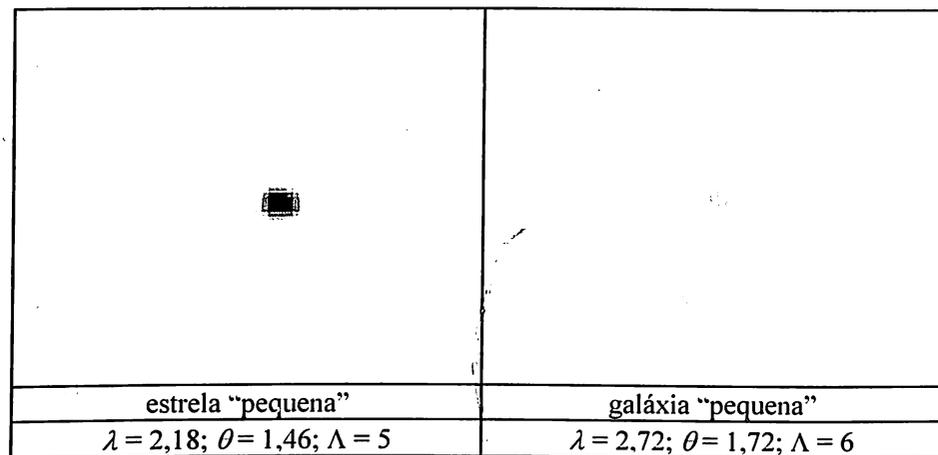


Figura 4 - Espectro de Padrões e atributos associados de 4 objetos utilizando a métrica do quarteirão

Considerações Finais

É fato comprovado que a multidisciplinaridade intrínseca na tecnologia que vem sendo desenvolvida ao longo do tempo para auxiliar estudiosos das profundezas do Universo, hoje auxilia o ser humano no seu dia-a-dia.

O princípio intuitivo da MM consiste em extrair informações a partir de transformações morfológicas utilizando elementos estruturantes cuidadosamente selecionados, tornando a MM atrativa para ser aplicada em muitos ramos da PDAI. A MM tem trazido grandes contribuições nas tarefas de Análise de Imagens, proporcionando resultados como: esqueletização de imagens, descrição de formas por granulometria, filtros morfológicos, extração de contornos, preenchimento de falhas nas imagens, entre outros. O propósito deste trabalho foi encontrar um compromisso no refinamento entre a detecção e as medidas dos objetos astronômicos a partir de uma interpretação inovadora utilizando ferramentas robustas da MM.

Os atributos de análise baseados no espectro de padrões mostraram-se adequados para a identificação de estrelas e galáxias, fortalecendo a idéia de que atributos morfológicos de fato são apropriados para o reconhecimento de padrões astronômicos, aproximando-se dos aspectos inerentes à classificação visual.

Como o intuito de obter resultados mais precisos, essa pesquisa continuará em busca de uma definição de uma família de elementos estruturantes mais apropriada para a granulometria, que possa levar em consideração informações de detalhes estruturais do objeto. Granulometria em níveis de cinza é um tema emergente e pouca literatura pode ser encontrada.

Os objetos mais fracos são comumente tratados de maneira subjetiva, requerendo, ainda, a assistência da visão especialista humana para identificar ou rejeitar tais objetos para a etapa de classificação. A intenção é implementar, de maneira totalmente automática, a abordagem morfológica associada a um classificador, utilizando as características de tamanho e forma extraídas das operações morfológicas.

Aliado ao objetivo principal de propor uma abordagem morfológica de representação dos objetos astronômicos, existiu, neste trabalho, uma idéia subjacente de promover o diálogo entre as diversas áreas de pesquisa, no nosso caso por meio da multidisciplinaridade entre PDAI, MM e Astronomia, visando reavivar o interesse da comunidade de PDAI pelas aplicações com imagens astronômicas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Candéas A.J., *Uma Abordagem Morfológica para a Caracterização de Objetos Estelares*, Tese de Mestrado, Universidade Federal de Pernambuco, Departamento de Informática, 1998.
2. Corwin, H.G., de Vaucouleurs, G. e de Vaucouleurs A., *Southern Galaxy Catalog*, University of Texas, Astronomy Department, 1985.
3. Humphreys, R.M. e Pennington, R.L., *Workshop on digitized optical sky surveys*, editado por C. Jaschek and H.T. MacGillivray, 1989.
4. Odewahn, S.C., Mumphreys, R.M., Aldering, G. e Thurmes, P., "Star-galaxy separation with a neural network - II - Multiple Schmidt plate fields", *Publications of the Astronomic Society of the Pacific*, 105:1354--1365, 1993.
5. Banon, G.J.F., "Characterization of translation-invariant elementary operators for gray-level morphology", *Neural, Morphological and Stochastic Methods in Image and Signal Processing*, 68--79, Proc. SPIE 2568, 1995.
6. Heijmans, H.J.A.M., *Morphological Image Operators*, Academic Press, Inc., 1994.
7. Serra, J., *Image Analysis and Mathematical Morphology, Vol. 2: Theoretical advances*, capítulo 5 - *Introduction to Morphological Filters*, editado por J Serra, Academic Press, Inc., 1988.
8. Vincent, L. e Soille, P., "Watersheds in digital spaces: an efficient algorithm based on immersion simulations", *IEEE Transactions on Pat. Anal. Mach. Intel.*, 13(6):583--598, 1991.
9. Facon, J., "Morfologia Matemática: Teoria e Exemplos", Editora Universitária Champagnat da Pontifícia Universidade Católica do Paraná, Curitiba, 1996.
10. Matheron, G., "Random Sets and Integral Geometry", Wiley, New York, 1975.
11. Maragos, P., "Pattern spectrum and multiscale shape representation", *IEEE Transactions on Pat. Anal. Mach. Intel.*, 11(7):701--716, 1989.
12. Sternberg, S.R., "Grayscale morphology", *Computer Vision, Graphics and Image Processing, Special edition on Mathematical Morphology*, 35:333--355, 1986.
13. Barrera, J., Banon, J.G.F. e Lotufo, R.A., "A Mathematical Morphology toolbox for the KHOROS system, Technical Report RT-MAC-9403, Instituto de Matemática e Estatística - Universidade Estadual de São Paulo, São Paulo, SP, Brasil, 1994.

POESIA PERNAMBUCANA MODERNA

(Breve Antologia)

(Organizada por César Leal)

Edição de Estudos Universitários
(Separata)
Recife, 1999