

UNIDADE 5 - Extração de atributos

Para reduzir a quantidade de informação ou para melhorar o desempenho do sistema na etapa de reconhecimento, é comum termos uma etapa chamada de extração de atributos, que irá transformar o segmento (ou segmentos) da imagem que interessa aos módulos de processamento posteriores em uma outra representação, geralmente um vetor de atributos. Os módulos posteriores, como o de reconhecimento de objetos, por exemplo, não irão trabalhar portanto diretamente na imagem, mas em um conjunto de valores extraídos da imagem. Existem centenas de técnicas de extração de atributos disponíveis e nas seções seguintes agrupamos essas técnicas em cinco grandes grupos: atributos baseados em cores, em forma, em textura, em pontos de interesse e em um último grupo de técnicas que permitem a combinação ou seleção de atributos variados. Atributos são também chamados por alguns autores de características e por outros de descritores.

5.1 Cores

O grupo mais básico ou elementar de atributos é formado por cores, incluídos aqui os valores de tons de cinza no caso de imagens “em preto e branco”. É possível trabalhar diretamente com os valores de todos os pixels que compõem o objeto ou cena de interesse ou então, como é mais comum, reduzir ou resumir estes valores através de alguma estatística, como a média, mediana e desvio padrão dos valores de cor que compõem a região de interesse. Também é muito comum o uso de histogramas de cores (apresentados na Seção 3.1, Página 26) que tem como grande vantagem poderem ser facilmente transformados em um vetor de atributos (representação muito utilizada na área da aprendizagem supervisionada). Embora a forma mais comum de se representar uma cor em meio digital seja através do sistema RGB, que indica com 3 números as quantidades de vermelho (*Red*), verde (*Green*) e azul (*Blue*) que compõem a cor de determinado pixel, existem outras opções, que dependendo do problema, podem ser mais interessantes. Chamamos de espaços de cores as diferentes formas de representação. Uma alternativa ao espaço de cores RGB que é também muito comum em sistemas de visão computacional, é o HSV ou Hue-Saturation-Value que em Português poderia ser

traduzido como Matiz-Saturação-Valor. O sistema HSV se aproxima mais da maneira como o ser humano interpreta as cores. O valor, por exemplo, está relacionado com a maneira com que classificamos uma cor como clara ou escura. Quanto maior o valor, mais clara é a cor. Já a saturação tem relação com a pureza ou vivacidade da cor, quanto maior a saturação, mais pura ou viva é a cor (ou seja, menos mistura de outras cores). O cinza é um exemplo de uma cor com valor de saturação bastante baixo, pois resulta da “mistura” de muitas cores. A matiz é aquilo que associamos aos principais nomes de cores, como azul, vermelho, amarelo, verde, roxo, etc. Os componentes H e S do sistema HSV estão mais associados à cromaticidade do próprio objeto enquanto o V depende muito mais da iluminação que está sendo aplicada (quanto mais iluminado o ambiente, maior o valor de V). É comum se descartar da análise o componente V quando se tem um sistema que precisa funcionar com diferentes intensidades de iluminação. O vídeo apresentado na Figura 32 explica através de imagens e usando os softwares ImageJ e GIMP o HSV, que também é conhecido como HSI.

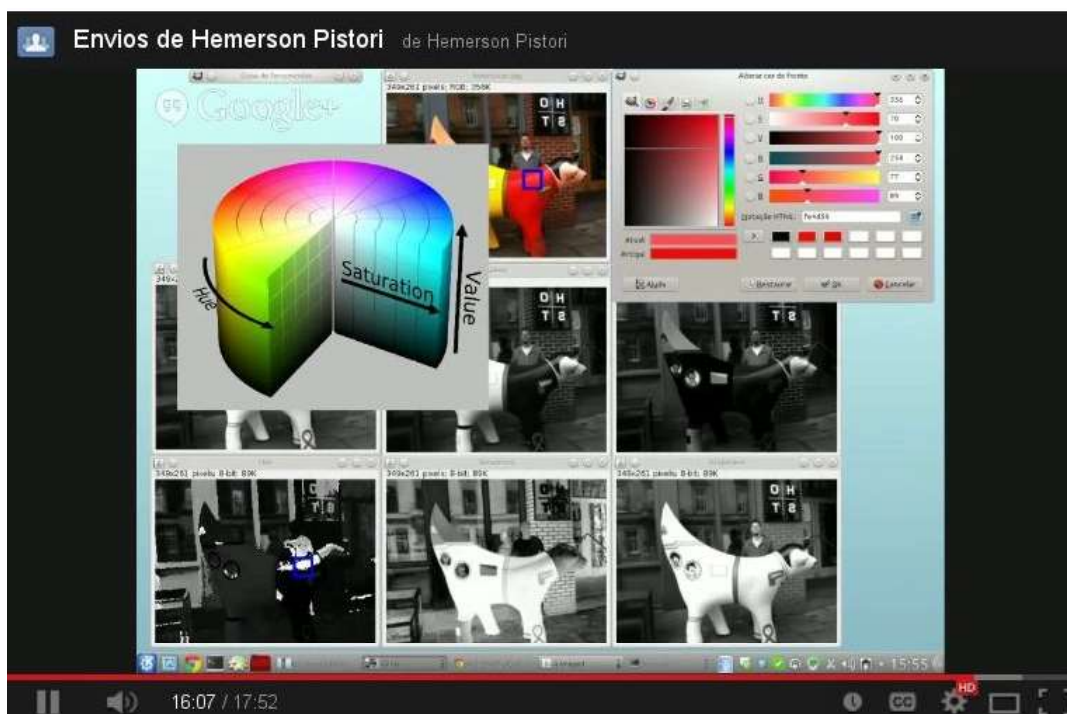


Figura 32: Vídeo sobre espaço de cores disponível em <http://www.youtube.com/watch?v=Chz9A-qVHI8&feature=share>

Na Figura 33, primeira linha, apresentamos 4 cores com variação apenas na matiz (Hue). Na segunda linha temos um exemplo de variação apenas na saturação (de uma saturação baixa, à esquerda, para uma saturação mais alta, à direita) e na última linha mostramos como fica uma variação apenas no valor (do mais baixo para o mais alto).

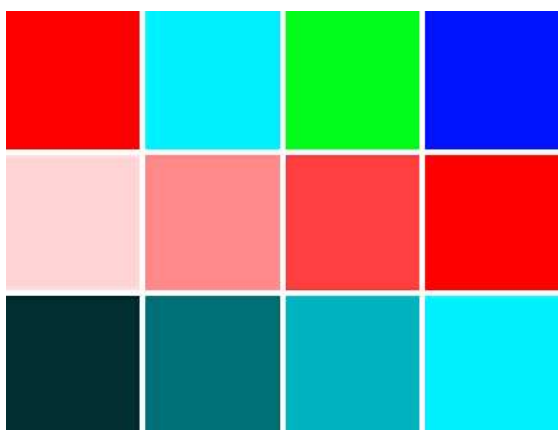


Figura 33: Cores com diferentes matizes (linha de cima), saturações (centro) e valores (linha de baixo)

5.2 Forma

Atributos de forma buscam caracterizar informações relacionadas com o contorno dos objetos. Existem, por exemplo, atributos de forma que estão associados à circularidade (quanto mais “circular” maior o valor do atributo) e ao alongamento (quanto mais longo e fino maior o valor) de um objeto, outros à topologia e ainda alguns geram histogramas que mantêm contagens sobre a ocorrência de curvas de variados ângulos ao percorrermos o contorno do objeto, como é o caso do algoritmo k-curvatura. Boa parte das técnicas utilizadas para se obter atributos de forma são baseadas em momentos estatísticos ou momentos da imagem.

5.3 Textura

O atributo de textura pode ser melhor entendido usando uma analogia com um conceito mais conhecido, e que inspira sua aplicação em visão computacional, que é a textura de tecidos. Diferentes tecidos possuem diferentes tramas que se repetem e que o caracterizam. Temos tecidos lisos, ou tecidos com listras horizontais, listras verticais, bolinhas, ou ainda imitando pele de onça, jacaré, enfim, o que caracteriza a textura é algum elemento que se repete várias vezes, lado a lado. Na Figura 34 temos 5 imagens que poderiam ser caracterizadas por tipos diferentes de textura. A Figura 35 apresenta as mesmas texturas agora em tons de cinza (sem as cores) para reforçar a importância da textura para se distinguir o conteúdo deste tipo de imagens. Atributos de textura podem ser também muito importantes na etapa de segmentação. Entre as técnicas mais utilizadas para caracterizar texturas numericamente (gerar vetores de atributos) estão as matrizes de co-ocorrência, os padrões binários locais e os bancos de filtros de Gabor.



Figura 34: Exemplos de diferentes tipos de textura retirados de imagens reais

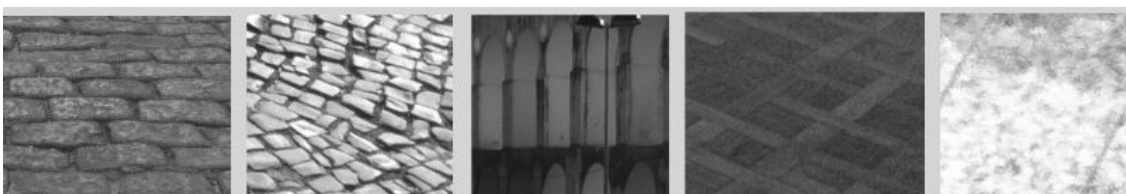


Figura 35: Mesmos exemplos da Figura 34 mas agora em tons de cinza

5.4 Pontos de interesse

Neste grupo de técnicas estamos incluindo algoritmos como o SIFT e o SURF, que se tornaram muito populares na área de visão computacional, recentemente, pelos bons resultados apresentados em diversos tipos de problemas de detecção de objetos. Uma das grandes vantagens apontadas para este tipo de técnica é a capacidade de detectar determinados tipos de objetos presentes em uma cena mesmo quando apenas partes destes objetos estão visíveis (chamamos este

problema de oclusão). É muito comum a utilização desse tipo de técnica juntamente com as técnicas de casamento de modelos que serão apresentada na Unidade 6. Tanto o SIFT quanto o SURF e outras técnicas similares funcionam em duas etapas principais. Na primeira etapa são identificados nas imagens os chamados pontos de interesse, que são regiões do objeto que espera-se que sejam mais representativas do mesmo e que na tarefa de encontrar estes objetos em diferentes situações possam ser mais úteis. Por exemplo, na Figura 36 temos a imagem de uma pomba à esquerda, ao centro marcamos com quadrados azuis e vermelhos 4 regiões do pombo. Supondo que o problema seja encontrar este pombo em outra imagens, as regiões marcadas em azul seriam mais úteis, pois carregam informações mais discriminativas em relação ao pombo. As regiões em vermelho são simplesmente regiões planas, uma clara e outra escura, que podem ocorrer em uma quantidade bem maior de outros tipos de objetos (não pombos) que o padrão que se vê nas regiões em azul. Ao lado direito da imagem mostramos as 4 regiões marcadas aumentadas para melhor visualização. O que diferencia as regiões mais interessantes das menos interessantes é geralmente a existência de mudanças bruscas nas cores, além de ângulos e curvas. A quantidade de regiões interessantes, ou pontos de interesse, varia de objeto para objeto, e existem diversas técnicas para encontrá-las, a maioria delas baseada na utilização de cálculo de derivadas direcionais ou gradientes.

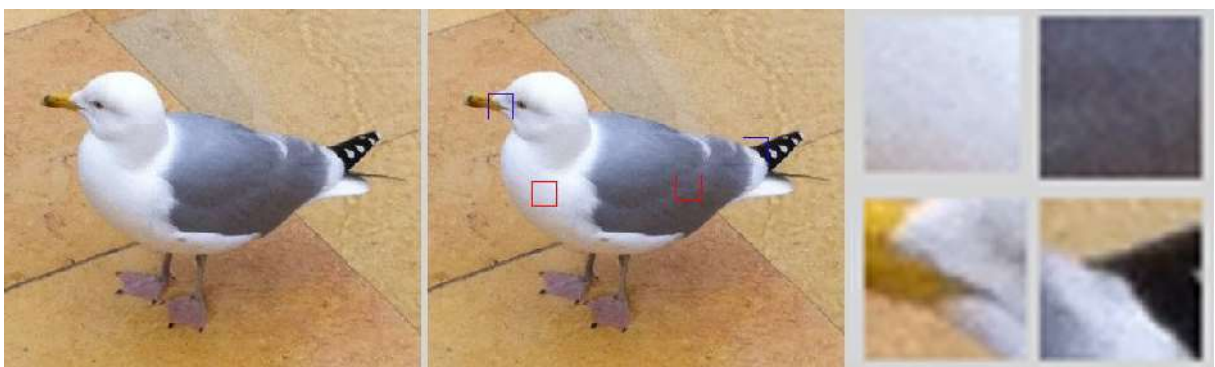


Figura 36: Possíveis pontos de interesse em uma imagem de pombo

Na segunda etapa, para cada ponto de interesse identificado, é gerado um vetor de atributos, também chamado de descritor do ponto de interesse. Esses descritores podem incluir atributos de cor, textura e forma, e são utilizados na busca por ocorrências do objeto em novas imagens.



Figura 37: Vídeo sobre o SURF disponível em <http://youtu.be/d5MbDgeoz3I>

Admitindo-se que basta encontrar

alguns dos descritores para concluirmos que o objeto está presente é que conseguimos fazer com que o sistema funcione mesmo em situação de oclusão parcial. Na Figura 37 temos a tela inicial de um vídeo que apresenta com mais detalhes uma das técnicas de extração de atributos baseadas em pontos de interesse, o SURF. Uma vez que o SURF e outras técnicas do mesmo grupo associam vários descritores (ou vetores de atributos) ao objeto que se pretende identificar, fica difícil a utilização das técnicas de aprendizagem supervisionada que precisam de um único vetor descrevendo cada objeto ou cena de interesse. Uma maneira de contornar esse problema é utilizando histogramas de palavras visuais (*Bag-of-Words* ou *Bag-of-Visual-Words*) que associam a cada objeto um histograma que conta o total de ocorrências de determinados tipos de descritores representativos. A escolha dos descritores representativos é realizada através de aprendizagem não-supervisionada, mais especificamente através do algoritmo *k*-médias (*k-means*). O vídeo representado na Figura 38 apresenta em detalhes o algoritmo *Bag-of-Words*.

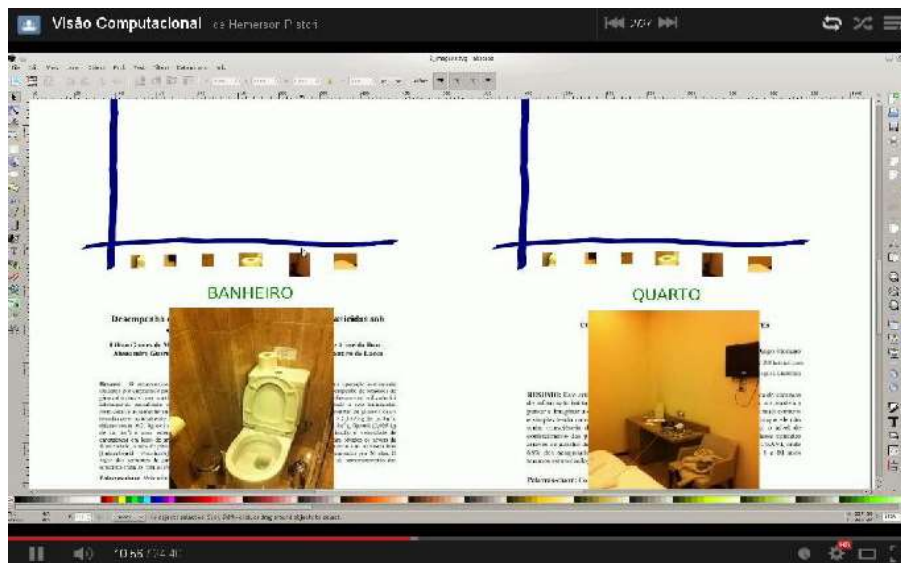


Figura 38: Vídeo sobre Bag-of-Words disponível em <http://youtu.be/J1OL6DxO5NM>

5.5 Seleção e Redução de atributos

Encontrar os atributos corretos para cada problema não é uma tarefa simples mas existem algumas técnicas que podem auxiliar. Uma estratégia é extrair o máximo de tipos de atributo possíveis usando uma variedade de algoritmos e deixar que um programa de computador escolha automaticamente um subconjunto do conjunto original de atributos que ofereça o melhor resultado em relação a um conjunto de imagens escolhidas para testar o sistema. Chamamos este tipo de estratégia de “seleção de atributos” e existem várias técnicas de busca estudadas na área da inteligência artificial, como a busca pelo melhor primeiro (*best-first search*), a busca por têmpera simulada (*simulated annealing*) e os algoritmos genéticos, que podem ser experimentadas (o software *weka* implementa algumas dessas técnicas). Uma outra forma para se chegar a um bom conjunto de atributos é através da redução de atributos, que diferentemente da seleção, realiza algum tipo de combinação dos atributos originais visando a obtenção de um conjunto geralmente menor que o original mas que possibilitem uma melhoria no desempenho ou a manutenção do desempenho original mas com uma quantidade menor de atributos.

Existem várias técnicas para combinar atributos sendo duas das mais utilizadas a análise discriminante de Fisher e a análise de componentes principais.