

Classificação de imagens de fotografias históricas usando detecção de linhas e paralelogramos pela presença de edifícios

Natália Cosse Batista

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
natalia@dcc.ufmg.br

Resumo

Neste trabalho é realizado o estudo de uma metodologia para classificar imagens de fotografias históricas baseada na detecção de edifícios. A metodologia envolve a detecção de bordas pelo algoritmo de Canny e a detecção de linhas e de paralelogramos usando a Transformada de Hough. Devido às características das imagens, são testados também alguns filtros para melhorar a qualidade com redução de ruído e ajuste de contraste. Resultados preliminares mostraram que os atributos usados na classificação possibilitam 62% de acerto, separando razoavelmente as classes edifícios da não-edifícios.

1. Introdução

A preservação de acervos de documentos históricos inclui, muitas vezes, a digitalização desses documentos, de maneira a evitar seu manuseio e conseqüentemente a aceleração do processo de degradação, produzindo centenas de gigabytes de informações. Essas grandes bases de dados necessitam de meios para catalogá-las e organizá-las provendo aos seus usuários - historiadores e pesquisadores, por exemplo - ferramentas para consulta e recuperação dos documentos desejados, de forma fácil e rápida.

As fotografias históricas representam uma parte significativa desse acervo, contendo os mais diversos temas. Apenas no Arquivo Público Mineiro ¹ são aproximadamente 80.000 fotografias tipicamente do século XIX e início do século XX, as quais 14.000 encontram-se preparadas para a digitalização e cerca de 6.000 já foram digitalizadas.

¹ Com um acervo de aproximadamente 1.400 metros lineares de documentos produzidos e acumulados por órgãos da Administração Pública de Minas Gerais e diversos arquivos privados, o Arquivo Público Mineiro recolhe e conserva importante patrimônio histórico e cultural, abrangendo os séculos XVIII, XIX e parte do século XX.

A anotação manual de um grande acervo é um processo caro e demorado [5], por isso fazem-se necessários sistemas que, além de auxiliarem a anotação, permitam consultas que não estejam restritas a ela. Visando melhorar a indexação e o acesso a esse material, este trabalho propõe a classificação automática das imagens de fotografias pela presença de edifícios, por meio de técnicas de Visão Computacional.

Trabalhos relacionados são discutidos na Seção 2. A metodologia proposta para esse trabalho é descrita na Seção 3. A Seção 4 detalha experimentos preliminares e a Seção 5 traz os resultados iniciais. Algumas conclusões e trabalhos futuros são relatados na Seção 6.

2. Revisão Bibliográfica

Os autores de [9] definem edifícios como objetos construídos pelo homem com contornos bem definidos e bordas retas. Essas estruturas contêm, na maioria das vezes, portas e janelas, formando um conjunto de paralelogramos alinhados. Ao passo que linhas extraídas de imagens de não-edifícios são geralmente distribuídas aleatoriamente. A detecção de fotografias que possuem fachadas de edifícios e de outras construções (igrejas, por exemplo), com o auxílio de técnicas de Recuperação de Imagens com Base no Conteúdo Visual (RIBC), pode ser usada na classificação de áreas entre urbanas ou rurais e na diferenciação entre ambientes externos e internos, colaborando com a indexação e recuperação dessas imagens.

As técnicas de RIBC são geralmente baseadas na análise de cores, formas e texturas das imagens. Tipicamente, a construção de um sistema RIBC compreende duas etapas: a representação das imagens e a comparação entre as representações das imagens da base com a imagem-consulta, para computação de um ranking de similaridade [15].

Segundo [10], as técnicas de representação mais usuais

podem ser divididas em duas grandes categorias: abordagens baseadas em visão e abordagens baseadas em modelos. As abordagens baseadas em visão, aspectos puramente quantitativos da imagem são computados, sendo os mais comuns os histogramas de cores e técnicas de análise de texturas. Por exemplo, as técnicas descritas em [4] e [13], que realizam a classificação de imagens de exteriores. Entretanto, conforme observado por [16], a classificação de imagens que contém estruturas construídas pelo homem não é satisfatória nestes trabalhos pois cor e textura não são informativas para esta classe.

Já as técnicas baseadas em modelos buscam a extração de modelos 3D dos objetos de interesse. Essas técnicas, oriundas principalmente dos trabalhos de visão computacional, requerem um conhecimento *a priori* da forma dos objetos. As técnicas baseadas em modelos são pouco aplicáveis a domínios amplos, pela imprevisibilidade do que pode ser encontrado em cada imagem, não só em termos de conteúdo, mas em termos de escala, ângulo de visão, etc. Em [9], por exemplo, é proposta a aplicação de princípios de agrupamento perceptivo, oriundos da psicologia Gestalt, para extração de elementos de nível intermediário que possam ser vistos como sugestões do conteúdo da imagem. Essa técnica começa pela extração de pequenos segmentos de reta que são, então, agrupados em retas maiores, junções em L, junções em U, grupos paralelos, grupos paralelos significativos e finalmente polígonos. As relações semânticas entre essas estruturas são usadas para detectar a presença de objetos construídos pelo homem, tais como edifícios, torres e pontes.

No trabalho em [6], Visão Computacional é utilizada para compreender cenas que em geral possuem grande diversidade de objetos, como cenas naturais contendo estradas, prédios, árvores, etc. O algoritmo proposto é baseado na detecção de ângulos e retas. Em testes anteriores executados com o programa fornecido pelos autores, nota-se que, em construções, os ângulos de janelas e portas são corretamente marcados, denotando certa organização geométrica, o que não ocorre nas imagens de paisagens naturais e plantações. Aliadas a esses resultados, informações de linhas e paralelogramos contidos nas imagens podem determinar a localização das portas e janelas nas imagens.

O algoritmo de detecção de paralelogramos descrito por [11] é capaz de detectar paralelogramos de tamanhos e orientações variadas mesmo em imagens ruidosas. Nesse trabalho, os autores fazem uso da Transformada de Hough para detectar paralelogramos. Além disso, esse método tem a vantagem de dividir a imagem em partições menores, formando uma grade retangular. Dessa forma, calcular a Transformada de Hough sobre cada partição da imagem reduz a complexidade computacional comparada à Transformada

de Hough global. Um algoritmo semelhante é também utilizado por [12] para detectar retângulos.

Com base nessas informações, a metodologia proposta neste trabalho será composta pela detecção de linhas utilizando a Transformada de Hough e detecção de paralelogramos usando o algoritmo de [11]. Adicionalmente, serão utilizados filtros para melhorar a qualidade das imagens, que possuem características bem distintas das imagens utilizadas pelos autores citados anteriormente. Nesses trabalhos, imagens sintéticas e fotografias recentes são usadas.

3. Metodologia proposta

A primeira etapa do trabalho é a seleção de imagens para os experimentos e sua separação em três classes: edifícios, intermediárias e não-edifícios (Figura 1). Essa separação foi sugerida em [9] e contém as seguintes características, adaptadas para esse trabalho:

- Edifícios: maior parte da imagem ocupada por fachada de edifício (parte externa). Pode conter pessoas, árvores, esculturas, carros e outros objetos ocupando pequena parte da imagem.
- Intermediárias: imagem contém alguma parte mínima de fachada, ou edifício distante. Pode conter vários objetos e pessoas ocupando maior porção da imagem.
- Não-edifícios: nenhuma parte da imagem contém edifícios.

O trabalho será realizado utilizando-se um subconjunto do acervo fotográfico digitalizado do APM de diversos temas (construções, paisagens, pessoas, carros, etc) e condições (iluminação, conservação, suporte fotográfico, ângulos, distância da cena, orientação, etc) contendo 90 imagens, 30 de cada classe. As fotografias têm aproximadamente 800x600 pixels em RGB com 24 bits (8 bits para o canal vermelho, 8 bits para o canal verde, 8 bits para azul) para a cor e formato JPEG.

A segunda etapa do trabalho é a detecção de bordas, realizada pelo algoritmo de Canny [2]. A Transformada de Hough será aplicada na imagem das bordas, e usada para a detecção de linhas e paralelogramos². No espaço de Hough, um ponto representa uma linha do plano da imagem e o valor do pico nesse ponto indica o tamanho da linha. O algoritmo de detecção de paralelogramos [11] parte da observação de [8] de que, no espaço de Hough, um paralelogramo é representado por quatro picos que guardam entre si algumas relações geométricas características (Figura 2):

2 O código-fonte do detector de paralelogramos foi fornecido pelos autores.



Figura 1. As duas primeiras colunas são exemplos de imagens da classe edifícios, as duas colunas centrais da classe intermediária e as duas últimas da classe não-edifícios.

- Esses picos aparecem aos pares, onde cada par tem o mesmo valor de θ , ou seja, tem o mesmo ângulo de inclinação;
- As intensidades dos dois picos são as mesmas, pois representam lados opostos do paralelogramo e têm portanto o mesmo tamanho;
- Os pares são separados no eixo θ do espaço de Hough por um valor α , onde α é um dos ângulos internos do paralelogramo.
- As distâncias verticais entre os picos no eixo ρ do espaço de Hough são as distâncias entre os lados do paralelogramo.

A terceira etapa do trabalho é a realização de testes preliminares com um conjunto de amostras para determinar parâmetros e limiares. As métricas escolhidas para os experimentos são:

1. Número total de linhas.
2. Número de linhas maiores ou igual a um limiar.
3. Razão entre as métricas 2 e 1, para efeito de normalização.
4. Número de paralelogramos.

O objetivo inicial é utilizar as métricas 3 e 4 para fazer a separação em classes.

A quarta etapa é a execução do experimento, com os parâmetros e limiares escolhidos na etapa anterior, no conjunto de 90 imagens. Inicialmente, a classificação das imagens será baseada nos limiares escolhidos pela análise

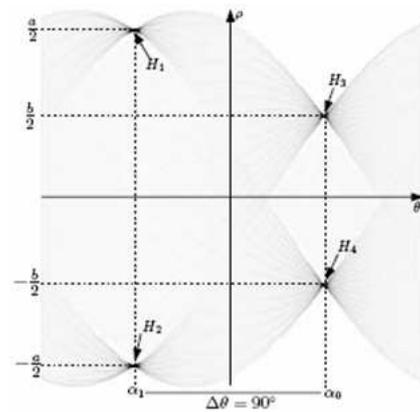


Figura 2. Transformada de Hough de um retângulo centrado na origem [12].

dos resultados preliminares. Como esta técnica é bastante limitada, pois é baseada em um espaço unidimensional em que os limiares escolhidos podem não descrever as diferenças entre as classes, será também utilizada a técnica de classificação por *Support Vector Machine* (SVM) [1]. Os resultados serão apresentados por meio de uma matriz de confusão, em que cada linha e cada coluna representam uma classe, cada elemento da matriz contém o número de imagens da classe dada pela linha classificadas como a classe dada pela coluna. A diagonal da matriz de confusão reflete o resultado da classificação para

as classes existentes. O esquema de todas as etapas de trabalho é mostrado na Figura 3.

4. Experimentos

4.1. Testes com imagens sintéticas

Testes preliminares³ foram realizados variando-se os parâmetros tamanho da partição (tile), lado mínimo do paralelogramo (RMin) e ângulo interno do paralelogramo em duas imagens das classes edifícios e não-edifícios. Observou-se que não foi detectado nenhum paralelogramo de ângulo interno 90° para RMin = 5 e RMin = 10 e tiles 10x10 e 15x15. Nos testes realizados com tiles 10x10, 15x15 e 30x30 e RMin = 5 e RMin = 10 não foram detectados paralelogramos correspondentes às janelas do edifício, mas vários outros que não coincidem com as bordas detectadas.

Esses resultados mostraram a necessidade de realizar experimentos com imagens ideais, sintéticas, com ruído mínimo, pois apenas no teste com tile = alt/2 x larg/2 e RMin = 5 alguns dos paralelogramos detectados estavam de acordo com as bordas oriundas das linhas formadas por portas, janelas e contornos do edifício, sugerindo uma melhor validação do detector para essa função.

Dois imagens sintéticas foram usadas para testar a influência dos parâmetros tamanho da partição (tile) e lado mínimo do paralelogramo (RMin). Os resultados mostraram que tiles próximos do tamanho das janelas da imagem sintética (aproximadamente 16x20 pixels) melhoram a detecção e que o RMin deve ser entre 5 ou 10 para essa imagem. Os testes foram realizados com o detector de paralelogramos configurado para encontrar paralelogramos genéricos. A Figura 4 mostra a tabela com os resultados. Pode-se concluir que os parâmetros que melhor diferenciam as imagens são tile 20x20 e RMin = 10. As imagens sintéticas e exemplos de resultados são encontrados na Figura 5.

4.2. Experimento com imagens de configurações distintas

Devido às características das imagens, foram testados dois filtros para melhorar a qualidade, o filtro gaussiano para redução de ruído e ajuste de contraste, ambos 3x3. O filtro gaussiano foi usado com $\sigma = 0,5$ e o filtro de ajuste de contraste com $\alpha = 0,2$, e correspondem aos filtros *gaussian* e *unsharp* do programa Matlab (Figura 6).

Também foram testadas uma redução de cor para aproximar a fotografia da imagem ideal, com poucos detalhes,

3 Todos os testes realizados neste trabalho e seus resultados são detalhados em <http://www.verlab.dcc.ufmg.br/doku.php?id=cursos:visao:2008-1:grupo08:index>.

Imagem sintética 1					
Tile	RMin	# linhas	#paralelogramos	# linhas >= 20	pico máximo
5x5	5	1675	5	0	12
10x10	5	586	0	100	20
15x15	5	368	10	113	30
20x20	5	339	33	129	40
20x20	10	245	22	129	40
20x20	15	123	0	112	40
20x20	20	90	0	90	40
30x30	5	177	10	129	60
40x40	5	136	16	115	80
alt/4xlarg/4	5	90	1	82	112
alt/2xlarg/2	5	8	4	8	160
Imagem sintética 2					
Tile	RMin	# linhas	#paralelogramos	# linhas >= 20	pico máximo
5x5	5	2416	2	0	13
10x10	5	1550	20	98	22
15x15	5	1004	71	2002	34
20x20	5	849	36	259	40
20x20	10	276	0	188	40
20x20	15	89	0	86	40
20x20	20	32	0	32	40
30x30	5	576	10	271	65
40x40	5	333	10	192	77
alt/4xlarg/4	5	144	4	109	116
alt/2xlarg/2	5	16	0	14	171

Figura 4. Resultados para as imagens sintéticas.

uma redução no tamanho das imagens para que o tempo de execução (por volta de 10 minutos por imagem) fosse menor, e a retirada das molduras que introduzem paralelogramos por causa de suas linhas e detalhes, por exemplo na Figura 7. Em [14] é proposta uma metodologia para a remoção automática de molduras da base de imagens do APM, que de fato afetam os resultados da recuperação e indexação de imagens com base no conteúdo.

O objetivo com estes novos testes é verificar o comportamento do detector de linhas e paralelogramos nas imagens com diferentes configurações e visualizar um possível limiar. Nesse experimento foram utilizadas 45 imagens, sendo 3 imagem para cada classe, nas 5 configurações diferentes:

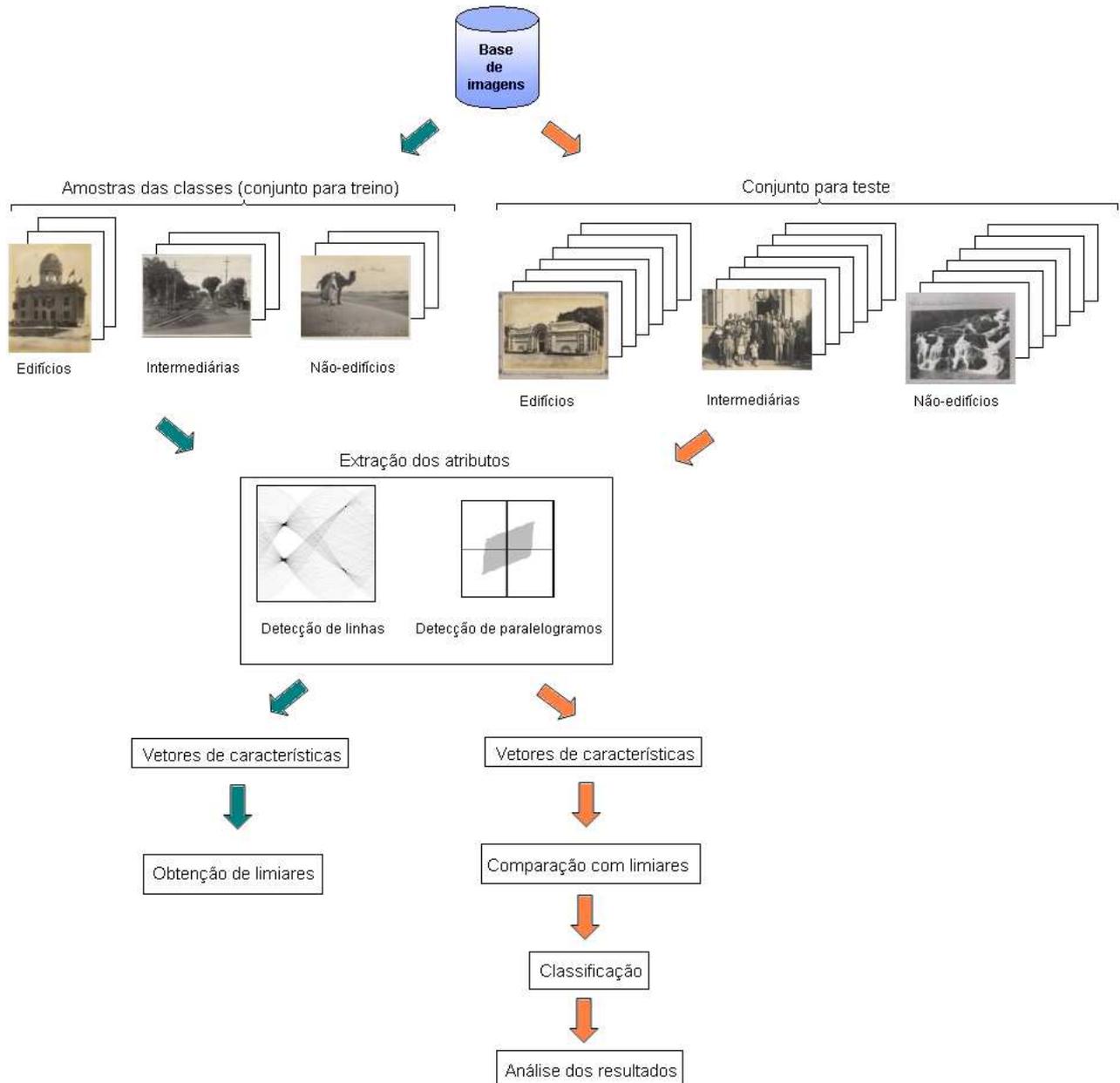


Figura 3. Resumo das etapas do trabalho.

1. Imagem original.
2. Imagem original reduzida em 50% do tamanho.
3. Imagem original reduzida em 50% sem a moldura (retirada manualmente).
4. Imagem original reduzida em 50% sem a moldura com apenas 3 níveis de cinza.
5. Imagem original reduzida em 50% sem a moldura com melhora do contraste e suavização.

O exemplo de uma imagem nas cinco configurações é mostrado na Figura 8.

Os resultados para tile 20x20 e RMin = 10 foram analisados para a obtenção de um limiar para as métricas (Figura 15). Entretanto, para a métrica 3 os resultados foram muito parecidos entre as classes e para os limiares 50, 60, 70, 90 e 100 os resultados foram iguais ou próximos de zero, portanto não sendo utilizada. Dessa forma, foram estabelecidos limiares apenas para o número de paralelogramos e o número de linhas para duas configurações, como mostrado

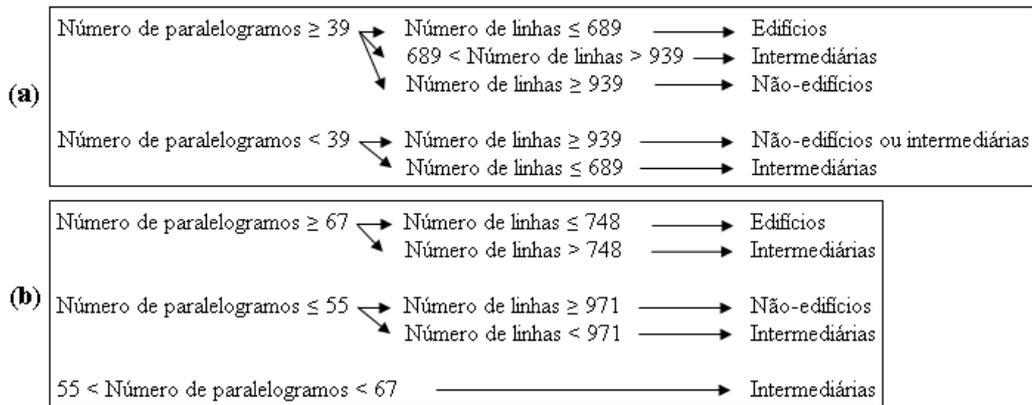


Figura 9. Limiares obtidos para as configurações (a) Imagem original reduzida em 50% sem a moldura e (b) Imagem original reduzida em 50% sem a moldura com melhora do contraste e suavização.

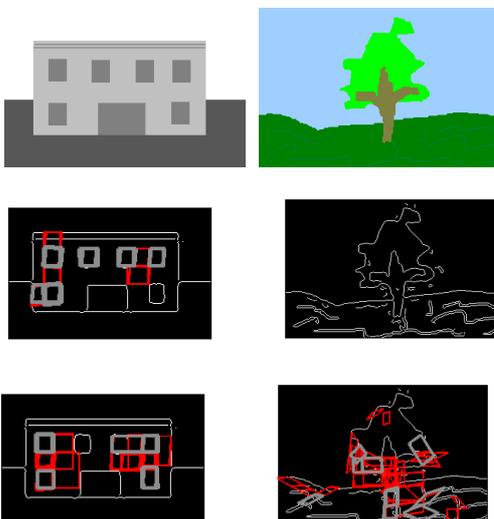


Figura 5. Imagens sintéticas (224 x 145 e 297 x 198, formato JPEG), resultados com tile 20x20 e RMin = 10, resultados com tile 30x30 e RMin = 5. Em cinza os paralelogramos detectados.

na Figura 9.

Os limiares encontrados foram usados para a classificação do conjunto de teste, com 90 imagens selecionadas aleatoriamente, 30 de cada classe. As imagens sofreram redução de resolução em 50% e as molduras presentes foram retiradas manualmente e para a segunda configuração, além disso, foram aplicados os filtros

gaussian =	0.0113	0.0838	0.0113
	0.0838	0.6193	0.0838
	0.0113	0.0838	0.0113
unsharp =	-0.1667	-0.6667	-0.1667
	-0.6667	4.3333	-0.6667
	-0.1667	-0.6667	-0.1667

Figura 6. Filtros 3x3 gaussiano e ajuste do contraste.

para melhorar a qualidade.

Os resultados são mostrados nas matrizes de confusão da Figura 10. Na matriz de confusão da configuração 3, os resultados foram 13%, 57% e 0% para as classes edifícios, intermediárias e não-edifícios, respectivamente, e na configuração 5, os resultados foram 1%, 83% e 0%. Aparentemente, os atributos usados na classificação possibilitam distinguir uma das classes, a intermediária, das outras duas, edifícios e não-edifícios. Entretanto, uma análise das matrizes de confusão revela que há uma concentração de valores na coluna da classe intermediária, o que significa que a maior parte das imagens foi classificada como intermediária e baixa concentração na diagonal, ou seja, a maioria das imagens não foi classificada corretamente. Isso indica que os atributos escolhidos tem valores próximos para todas as classes e que os atributos e/ou limiares escolhidos para a classificação não são suficientes para distinguir as três classes.



Figura 7. Imagem das bordas com paralelogramos detectados.

4.3. Experimento com SVM

SVM é uma técnica para classificação de dados, a qual, dado um conjunto de treino em que cada instância contém o rótulo de uma classe e determinadas características, produz um modelo que prevê a classe das instâncias de dados de um conjunto de teste.

Os vetores de treino são mapeados para um espaço de dimensão mais alta, de acordo com a função kernel utilizada. Então SVM procura um hiperplano linear ótimo com um margem máxima nesse espaço dimensional, permitindo a separação das classes com erro mínimo (Figura 11). Existem vários tipos de kernel, como o linear, polinomial, gaussiano (RBF: radial basis function) e sigmoidal. Nesse trabalho será usado o kernel RBF, que possui parâmetros C e γ . $C > 0$ indica o limite entre o erro de treinamento e a margem e γ é um coeficiente intrínseco do kernel que determina a largura do vetor de suporte.

Apesar de que em algumas situações o kernel RBF não se aplica (por exemplo quando o número de características é muito grande, mas não é o caso deste trabalho), no trabalho [7] são apontadas algumas de suas vantagens:

- O kernel RBF realiza um mapeamento não-linear das amostras para um espaço de dimensão mais alta, o que é importante quando a relação entre os rótulos das classes e as características é não-linear.
- Além da vantagem acima em relação ao kernel linear, que é um caso especial do RBF, possui o mesmo desempenho que o kernel linear para determinados parâmetros (C, γ).
- Como o número de hiperplanos influencia na complexidade da seleção do modelo, em relação ao kernel polinomial, o kernel RBF possui menos hiperparâmetros.

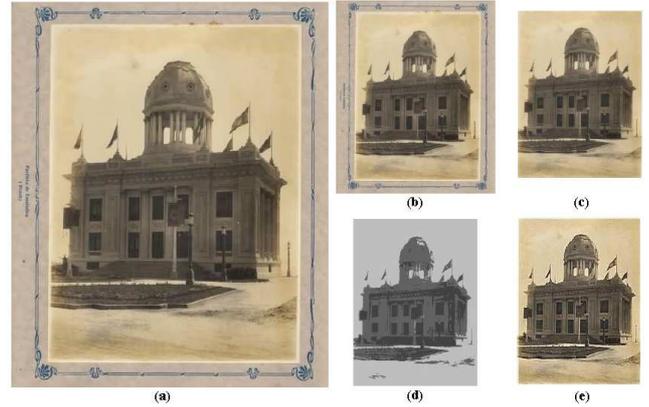


Figura 8. Exemplos das configurações testadas. (a) Imagem original. (b) Imagem original reduzida em 50% do tamanho. (c) Imagem original reduzida em 50% com a moldura. (d) Imagem original reduzida em 50% com a moldura com 3 níveis de cinza. (e) Imagem original reduzida em 50% com a moldura com melhora do contraste e suavização.

- O kernel RBF possui menos dificuldade numéricas.

Novos testes informais foram realizados para determinar parâmetros e características que pudessem obter melhores resultados utilizando SVM [3]. Foram escolhidas 45 imagens de treino e 45 imagens de teste, 15 de cada classe e 9 características extraídas com tile 50x50 e $R_{Min} = 10$:

- número de paralelogramos
- número de linhas
- número de paralelogramos / número de linhas
- número de linhas ≥ 40 / número de linhas
- número de linhas ≥ 50 / número de linhas
- número de linhas ≥ 60 / número de linhas
- número de linhas ≥ 70 / número de linhas
- número de linhas ≥ 80 / número de linhas
- número de linhas ≥ 100 / número de linhas

Os resultados estão na Figura 12.

5. Análise dos resultados

Os resultados obtidos com SVM foram consideravelmente melhores do que por classificação simples por limiar, sendo 62,2% (28 imagens classificadas corretamente em 45) e 31,1% (28 imagens classificadas corretamente em 90, no melhor caso), respectivamente. Na matriz

Classes	1	2	3	4
1	4	14	6	6
2	3	17	2	8
3	2	24	0	4

Classes	1	2	3
1	3	27	0
2	2	25	3
3	0	30	0

Figura 10. Matriz de confusão para a configuração 3 e para a configuração 5. Classe 1 = edifícios, Classe 2 = intermediárias, Classe 3 = não-edifícios e Classe 4 = não foi possível classificar.

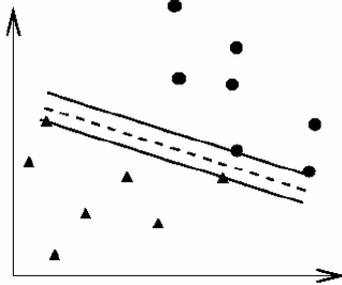


Figura 11. Problema de classificação binário. As classes são representadas por triângulos e círculos.

de confusão dos resultados da classificação com SVM há maior concentração de valores na diagonal, porém a classe edifícios confunde-se com a intermediária. Isso deve-se ao fato de que na classe intermediária as imagens podem conter estruturas, mesmo que em menor proporção. Entretanto a classe edifícios diferencia-se da não-edifícios de forma satisfatória.

Para efeito de comparação, a Figura 13 traz os resultados do trabalho de Iqbal [9]. Deve-se notar que no trabalho citado, a base de imagens utilizada foi obtida diretamente de câmera digital em cidades cujos edifícios são na maior parte modernos, com maior quantidade de andares, janelas e linhas retas, e são imagens menos ruidosas que a base do APM (Figura 14). Como não foi possível obter tal base, a comparação será restrita aos dados coletados do artigo dos autores.

Analisando-se a matriz de confusão do trabalho de Iqbal, nota-se que 35,4% das imagens intermediárias foram clas-

	1	2	3	
1	7	7	1	Classe 1 = 46,67 %
2	2	11	2	Classe 2 = 73,33 %
3	2	3	10	Classe 3 = 66,67 %

Figura 12. Matriz de confusão e porcentagem de acertos da classificação com SVM para as 45 imagens de teste.

Classe	Iqbal	Este trabalho
Edifícios	77,65%	46,67%
Intermediárias	51,04%	73,33%
Não-edifícios	81,43%	66,67%

Figura 13. Porcentagem de imagens classificadas corretamente no trabalho de [9] e neste trabalho.

sificadas como edifícios, entretanto apresenta resultados satisfatórios para as classes edifícios e não-edifícios. A porcentagem de acerto total foi de 73,5% (361 imagens classificadas corretamente em 491).

6. Conclusões e trabalhos futuros

Embora o algoritmo de detecção de paralelogramos seja capaz de detectar paralelogramos de tamanhos e orientações variadas mesmo em imagens ruidosas, observou-se que os paralelogramos detectados nas imagens não coincidem exatamente com os contornos das janelas e portas dos



Figura 14. Exemplos de imagens da base utilizada no trabalho [9].

edifícios, como esperado. Um ajuste melhor dos parâmetros tile e RMin poderia levar a resultados mais satisfatórios nessas imagens, levando-se em conta uma média do tamanho das janelas, por exemplo.

A características escolhidas possibilitam 62% de acerto na classificação, porcentagem que pode ser melhorada incluindo-se mais características no modelo, principalmente levando-se em conta o posicionamento dos paralelogramos, que parecem mais alinhados nas imagens de edifícios do que das imagens de não-edifícios.

Como trabalhos futuros, tem-se o objetivo de: utilizar um conjunto de imagens maior para o treinamento e testes, selecionar novas características, por exemplo um histograma de orientações dos paralelogramos, uso de métodos quantitativos e estatística para caracterizar o conjunto de imagens, projetar e comparar os experimentos de forma mais objetiva, testar outros filtros para melhorar qualidade das imagens.

Referências

- [1] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [2] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
- [3] C.-C. Chang and C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001. Software disponível em <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [4] X. Feng, C. K. I. Williams, and S. N. Felderhof. Combining belief networks and neural networks for scene segmentation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24, pages 467–483, 2002.
- [5] T. Gevers and A. W. M. Smeulders. *Content-Based Image Retrieval: An Overview*, chapter 8. IMSC Press Multimedia Series. Prentice Hall, 1st edition, July 2004.
- [6] C. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [7] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. *A practical guide to support vector classification*. Taiwan, May 2008. Published online.
- [8] D. Ioannou and E. Dugan. Parallelogram detection in a digital image with the use of the hough transform. In *Proceedings of the IEEE Computer Society International Conference on Pattern Recognition*, volume II, pages 532–536. IEEE Computer Society, 1996.
- [9] Q. Iqbal and J. Aggarwal. Retrieval by classification of images containing large manmade objects using perceptual grouping. *Pattern Recognition*, 35(7):1463–1479, July 2002.
- [10] Q. Iqbal and J. K. Aggarwal. Feature integration, multi-image queries and relevance feedback in image retrieval. In *Proceedings of 6th International Conference on Visual Information Systems*, pages 467–474, Sep. 24–26 2003.
- [11] C. Jung and S. Rosito. Parallelogram detection using the tiled hough transform. In *Proceedings of 13th International Conference on Systems, Signals and Image Processing*, volume I, pages 177–180, 2006.
- [12] C. R. Jung and R. Schramm. Rectangle detection based on a windowed hough transform. In *Proceedings of XVII Brazilian Symposium on Computer Graphics and Image Processing*, pages 113–120. IEEE Computer Society, 2004.
- [13] S. Konishi and A. L. Yuille. Statistical cues for domain specific image segmentation with performance analysis. In *Proceedings of the 2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume I, pages 135–132, 2000.
- [14] A. P. B. Lopes, D. L. Flam, N. C. Batista, S. E. F. de Avila, J. M. Almeida, and A. de A. Araújo. Frame detection in historical photographs using tiled hough transform. 2008. Submetido para publicação.
- [15] N. Sebe, Q. Tian, E. Louprias, M. Lew, and T. Huang. Evaluation of salient point techniques. *Image and Vision Computing*, 21(13-14):1087–1095, December 2003.
- [16] A. Vailaya, A. K. Jain, and H. J. Zhang. On image classification: City images vs. landscapes. *Pattern Recognition*, 31:1921–1936, 1998.

	Imagens	Paralelogramos	# linhas	# linhas >= 20	Razão 20	# linhas >= 30	Razão 30	# linha >= 40	Razão 40	Pico máx
11	AB-05-3-012.jpg	280	3290	2886	0,877	1480	0,450	474	0,144	48
11	AB-05-3-015.jpg	176	3060	2723	0,890	1636	0,535	639	0,209	56
11	AB-05-3-033.jpg	267	3792	3333	0,879	1713	0,452	587	0,155	49
12	SI-51(30).jpg	87	3910	3056	0,782	1514	0,387	534	0,137	45
12	TG-1-050.jpg	23	2615	2217	0,848	1132	0,433	443	0,169	47
12	TG-2-031.jpg	259	3182	2678	0,842	1041	0,327	165	0,052	47
13	AB-06-1-027.jpg	55	4227	3326	0,787	1115	0,264	288	0,068	48
13	AB-08-012(04).jpg	259	4188	3356	0,801	911	0,218	40	0,010	44
13	JP-11-2-043.jpg	119	4442	3648	0,821	1207	0,272	310	0,070	47
21	AB-05-3-012_50.jpg	115	959	859	0,896	416	0,434	151	0,157	43
21	AB-05-3-015_50.jpg	98	900	826	0,918	502	0,558	188	0,209	47
21	AB-05-3-033_50.jpg	76	1021	809	0,792	403	0,395	177	0,173	46
22	SI-51(30)_50.jpg	18	1332	1096	0,823	552	0,414	178	0,134	60
22	TG-1-050_50.jpg	25	814	686	0,843	346	0,425	97	0,119	40
22	TG-2-031_50.jpg	102	862	680	0,789	239	0,277	30	0,035	48
23	AB-06-1-027_50.jpg	37	1033	877	0,849	332	0,321	90	0,087	41
23	AB-08-012(04)_50.jpg	42	1343	1166	0,868	474	0,353	56	0,042	50
23	JP-11-2-043_50.jpg	13	1091	880	0,807	319	0,292	111	0,102	45
31	AB-05-3-012_sm3c50.jpg	31	548	464	0,847	125	0,228	12	0,022	44
31	AB-05-3-015_sm3c50.jpg	126	688	598	0,869	300	0,436	53	0,077	59
31	AB-05-3-033_sm3c50.jpg	65	581	443	0,762	154	0,265	23	0,040	47
32	SI-51(30)_sm3c50.jpg	5	839	630	0,751	268	0,319	79	0,094	71
32	TG-1-050_sm3c50.jpg	12	561	464	0,827	174	0,310	14	0,025	44
32	TG-2-031_sm3c50.jpg	83	723	590	0,816	216	0,299	12	0,017	43
33	AB-06-1-027_sm3c50.jpg	1	548	445	0,812	150	0,274	23	0,042	42
33	AB-08-012(04)_sm3c50.jpg	37	1212	1038	0,856	423	0,349	50	0,041	52
33	JP-11-2-043_sm3c50.jpg	20	817	625	0,765	134	0,164	3	0,004	40
41	AB-05-3-012_sm50.jpg	154	630	550	0,873	185	0,294	21	0,033	47
41	AB-05-3-015_sm50.jpg	148	689	600	0,871	302	0,438	58	0,084	60
41	AB-05-3-033_sm50.jpg	39	653	492	0,753	165	0,253	31	0,047	44
42	SI-51(30)_sm50.jpg	20	1221	953	0,781	480	0,393	139	0,114	68
42	TG-1-050_sm50.jpg	25	660	540	0,818	221	0,335	24	0,036	43
42	TG-2-031_sm50.jpg	66	734	603	0,822	198	0,270	18	0,025	45
43	AB-06-1-027_sm50.jpg	33	1233	1014	0,822	378	0,307	34	0,028	43
43	AB-08-012(04)_sm50.jpg	42	1343	1166	0,868	474	0,353	56	0,042	50
43	JP-11-2-043_sm50.jpg	22	939	722	0,769	201	0,214	5	0,005	40
51	AB-05-3-012_sm50GU.jpg	195	724	675	0,932	291	0,402	50	0,069	56
51	AB-05-3-015_sm50GU.jpg	90	748	670	0,896	385	0,515	92	0,123	60
51	AB-05-3-033_sm50GU.jpg	67	733	591	0,806	256	0,349	55	0,075	60
52	SI-51(30)_sm50GU.jpg	59	1289	1043	0,809	560	0,434	172	0,133	72
52	TG-1-050_sm50GU.jpg	17	658	569	0,865	283	0,430	44	0,067	46
52	TG-2-031_sm50GU.jpg	83	783	683	0,872	304	0,388	57	0,073	55
53	AB-06-1-027_sm50GU.jpg	55	1209	1054	0,872	471	0,390	103	0,085	49
53	AB-08-012(04)_sm50GU.jpg	30	1402	1273	0,908	657	0,469	136	0,097	54
53	JP-11-2-043_sm50GU.jpg	26	971	827	0,852	323	0,333	37	0,038	44

Figura 15. Resultados para as imagens usadas como amostras. As cores (exceto amarelo) indicam possibilidade de determinar um limiar. Na numeração da primeira coluna, o primeiro algarismo indica a configuração e o segundo, a classe da imagem. Os limiares 50, 60, 70, 90 e 100 não foram listados pois seus resultados foram iguais ou próximos de zero.