Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Computação

USO DE FERRAMENTAS DE MORFOLOGIA MATEMÁTICA PARA SEGMENTAÇÃO DE IMAGENS DE POPULAÇÃO EM AMBIENTE RESTRITO FECHADO

Relatório de atividades apresentado pelo bolsista **FRANSÉRGIO EVARISTO VIEIRA,** sob orientação de **JANDER MOREIRA**, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Processo n º 680067/94-9

São Carlos
1996

SUMÁRIO

Projeto:	Uso de ferramentas de Morfologia Matemática para segmentação de imagens de população em ambiente restrito fechado.
Orientador:	Prof. Jander Moreira
Orientado:	Fransérgio Evaristo Vieira
Instituição:	Departamento de Computação - UFSCar
Tipo de bolsa:	ITI (PROTEM-CC - Fase II)
Período:	1 de abril de 1995 a 31 de março de 1996 (12 meses)
Processo:	(AnIMoMat) - 680067/94-9

1 INTRODUÇÃO

O monitoramento em tempo real de áreas públicas com alta densidade de pedestres normalmente é executado com limitações por observadores humanos. Estas limitações consistem no curto período de observação e na variação de confiabilidade [VELA93]. Este controle ou monitoramento pode ser efetuado com o intuito de melhorar o projeto de locais públicos, avaliar procedimentos de conduta e segurança pública, reduzir custos operacionais etc. Estudos quantitativos sobre movimentos de pedestres são eficientes quando a densidade de pedestres é baixa. Estes métodos não são indicados para situações de alta densidade de pedestres. São consideradas situações de alta densidade de pedestres aquelas com densidade maior que 1,25 pedestres/m^2 [POLU83]. Entretanto, é essencial a investigação de técnicas de aquisição automáticas para situações de alta densidade de pedestres, o que pode levar ao desenvolvimento de modelos mais realísticos para planejamento e projeto de áreas públicas. Também seria possível, com a utilização de sistemas *on-line*, a detecção de problemas na movimentação dos pedestres momentos antes e depois que eles ocorram [VELA83].

1.1 MORFOLOGIA MATEMÁTICA

A palavra morfologia vem do grego e significa estudo das formas. O nome Morfologia Matemática, hoje em dia consagrado, apareceu na época no campus de Fontainebleau na França. Consistente com o significado literal, o propósito original era analisar estruturas geométricas em imagens microscópicas de amostras de rochas ou metais e relacionar os resultados com propriedades físicas dos materiais. Sob um ponto de vista prático, esta técnica tem aplicações em vários problemas de análise de imagens (e.g. restauração, segmentação, medida, descrição simbólica etc.).

Fundada por Jean Serra e Georges Matheron nos anos sessenta [SERR82], a Morfologia Matemática tornou-se uma disciplina muito importante em Processamento de Sinais e Imagens. Seu domínio de atuação cobre um largo espectro de aplicações em Processamento de Imagens que vai desde o tratamento de imagens microscópicas até o processamento de imagens de satélites, aplicações industriais, médicas, automação de escritório etc. Como teoria matemática, ela se intersecciona com uma família de outras teorias matemáticas: probabilidade, estatística, topologia, teoria dos reticulados, lógica nebulosa e teoria da medida. Os operadores clássicos aplicados a funções reais são os operadores lineares invariantes por translação (i.t.). Estes operadores possuem uma característica singular: existe uma relação um para um entre o conjunto de operadores lineares i.t. e o conjunto das funções reais. A Morfologia Matemática estuda decomposições de operadores (i.e. mapeamentos ou transformações) entre reticulados completos em termos de quatro classes de operadores elementares: dilatações, erosões, anti-dilatações e anti-erosões [BANO94].

Hoje, a Morfologia Matemática é uma matéria intensivamente pesquisada em universidades e centros de pesquisa de todo o mundo, contando com dois congressos internacionais específicos sobre o tema.

2 OBJETIVOS

O objetivo principal é a segmentação de imagens com pedestres em um ambiente controlado, utilizando operadores da Morfologia Matemática para avaliar o número de pessoas presentes na cena. Como objetivo secundário enquadra-se a aplicabilidade dos operadores no problema sob pontos de vistas distintos. Com a possibilidade de aquisição de imagens coloridas, uma comparação de desempenho entre os resultados para o caso monocromático e colorido deve também ser considerado.

3 RECURSOS UTILIZADOS

O trabalho foi desenvolvido utilizando-se uma estação de trabalho da SUN Microsystems (SPARCstation-2), cujas características compreendem grande poder de processamento, grande capacidade de memória secundária (discos rígidos) e capacidade gráfica de alta resolução, usando sistema operacional padrão UNIX (SunOS). A interface gráfica opera no padrão X11, usando gerenciador de janela próprio (OpenLook).

O software que deu suporte ao desenvolvimento do procedimento de segmentação e análise de resultados foi o KHOROS System (X Window e UNIX). Foi utilizada a interface de programação visual CANTATA do KHOROS. Particularmente quanto às ferramentas morfológicas, foi utilizada a caixa de ferramentas de Morfologia Matemática MMach.

Para obtenção das imagens que foram filmadas no próprio campus da Universidade Federal de São Carlos, foi utilizado uma câmera filmadora Panasonic (VHS) apoiada sobre um suporte. Trata-se de uma filmadora comum sem nenhuma característica profissional. Para digitalizar as imagens foi utilizada uma placa digitalizadora TARGA ligada a um microcomputador pentium. As imagens obtidas foram armazenadas no formato tga para utilização no trabalho.

Segue uma descrição mais detalhada das imagens testes:

Tipo: RGB TRUE COLOR (24 bits)
Largura: 511 pixels
Altura: 486 pixels
Resolução: 96 pixels/polegada

3.1 O SISTEMA KHOROS

KHOROS é um software projetado para ambientes de pesquisa em processamento de imagens. Criado no Departamento de Engenharia Elétrica e Engenharia de Computação na Universidade do Novo México em Albuquerque, EUA, tem se tornado muito popular. Uma das características mais poderosas do Khoros é o **CANTATA**, sua interface com alto grau de abstração. Cantata é um programa que proporciona uma linguagem orientada a fluxo de dados expressada graficamente, a qual provê uma interface de programação visual para o ambiente do Khoros. Um programa no Cantata é chamado de *workspace*. Os programas dos usuários podem ser organizados em subsistemas independentes chamados de *toolboxes* (caixas de ferramentas), sendo integrados facilmente ao sistema.

3.2 A CAIXA DE FERRAMENTAS DE MORFOLOGIA MATEMÁTICA MMACH

A caixa de ferramentas de morfologia matemática MMach que foi desenvolvida em conjunto por FEE-UNICAMP, IME-USP e INPE. Possui ferramentas para trabalhar com sinais de uma dimensão, imagens binárias e em níveis de cinza. Cada família de operadores morfológicos é apresentado como um sub-menu do menu principal da caixa de ferramentas. Segundo a teoria de Morfologia Matemática, a maioria dos operadores são construídos pela composição de operações e operadores elementares, e a toolbox implementa estes princípios.

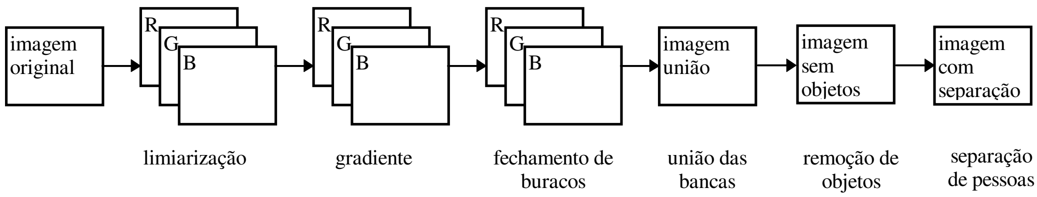


Figura 1. Esquema simplificado do fluxo de processamento. Omite-se a rotulação final.

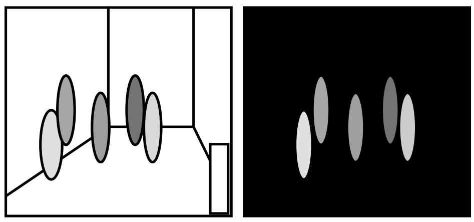


Figura 2. Remoção do fundo, deixando apenas os objetos presentes.

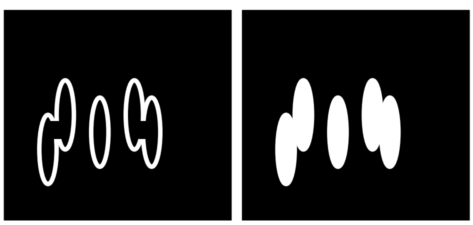


Figura 3. Contornos extraídos e fechamento de buracos (referências à ilustração da Figura 2).

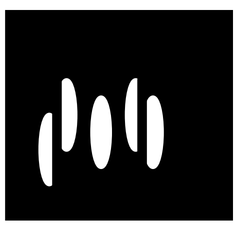


Figura 4. Separação de objetos próximos (com referência aos esquemas da Figura 2 e da Figura 3).

4 METODOLOGIA

A forma de tratamento que se propõe é esquematizada como um fluxo de procedimentos específicos, cada qual gerando resultados intermediários, até que se possa obter como resultado final o número de pessoas presentes na imagem. Um esquema simplificado do fluxo é dado na Figura 1.

Dado que se tem certo controle sobre o ambiente, é possível fazer a aquisição de uma imagem de referência, com a presença somente dos objetos fixos, ou seja, sem os pedestres. Fica viável, então, a eliminação do fundo das imagens a serem segmentadas, a partir da "remoção" da parte relativa à imagem de referência. Este processo é obtido pela diferença absoluta entre as duas imagens em questão. Como primeiro resultado intermediário, passa-se a dispor da imagem com somente as pessoas presentes, removido o ambiente (vide Figura 2). Este resultado, devido às diferenças de aquisição, nunca é perfeito, deixando pequenas regiões também na área considerada como fundo, o que é eliminado no processamento posterior.

Considerando que a imagem é colorida, a abordagem foi a extração das três bandas existentes para processamento individual (como esquematizado na Figura 1). Como cada banda da imagem é equivalente uma imagem em níveis de cinza, no que se refere a sua representação, é aplicada uma limiarização produzindo imagens binárias (preto e branco), as quais serão tratadas pelos operadores morfológicos.

O próximo passo, então, é a utilização do operador gradiente morfológico para extração das bordas das pessoas, obtendo-se os contornos das regiões na imagem que as caracterizam. Ao resultado da extração de bordas, aplica-se o operador de fechamento de buracos (*close holes*), sendo os contornos preenchidos em suas áreas internas (veja-se a Figura 3). Isto equivale a segmentar, em cada banda, áreas fechadas correspondentes às pessoas. Neste ponto as três imagens binárias correspondentes ao tratamento de cada banda são combinadas por operadores de união. Este resultado intermediário complementa as eventuais falhas de segmentação de cada banda em particular, obtendo-se um resultado melhor que se o tratamento fosse aplicado a uma imagem monocromática.

Pequenos objetos podem, ainda, estar presentes na imagem, como resquícios da segmentação em processo. Sendo assim, é necessário um pós-processamento para corrigir estas ocorrências. Para isto, um operador de N-erosão e um operador de N-dilatação são aplicados em conjunto, formando o que é conhecido como um filtro morfológico.

Quando, porém, as pessoas se encontram perto umas das outras, dois ou mais pedestres são detectados como um único objeto. Considerando que as pessoas presentes caracterizam-se como formas verticalmente dispostas, com relação entre altura e largura típicas, tal elemento estruturante utiliza este fato para remover conexões entre pessoas próximas. Posteriormente ao filtro, então, um operador de N-abertura com um elemento estruturante alongado na vertical é aplicado. Um esquema desta separação é apresentado na Figura 4.

Finalmente, uma rotulação na qual cada objeto (pessoa) presente na imagem recebe um rótulo distinto é efetuada. Avaliando, assim, o número de objetos detectados estima-se o número de pessoas presentes na imagem.

5 RESULTADOS E DISCUSSÃO

O tratamento consiste em submeter as imagens capturadas a uma série de operações morfológicas, utilizando a *toolbox* MMach. Através das operações morfológicas é possível manipular o conteúdo da imagem de forma a extrair as informações de interesse. No caso deste trabalho, em particular, as informações desejadas consistem nas pessoas presentes, considerando avaliar qual o número delas na cena.

Em sua grande parte, metodologias para processamento digital de imagens operam sobre imagens binárias (em dois níveis: preto e branco) ou as chamadas monocromáticas (em diferentes tons de cinza). Tratamentos especificamente projetados para aplicação a imagens coloridas são, de certa forma, recentes e têm aplicabilidade limitada em relação à Morfologia Matemática, em virtude da maior parte da literatura indicar imagens binárias e em níveis de cinza. [GONZ92]

No projeto desenvolvido foi possível dispor-se de imagens coloridas RGB (denominadas comumente *true color*, com 2^{24} possíveis cores). Desta forma, procurou-se adaptar as idéias iniciais de abordagem do problema de segmentação de modo a aproveitar a maior quantidade de informação disponível[1]. Assim, o projeto passou a compreender o tratamento de imagens coloridas.

Como primeiro passo, é processada uma ***remoção do fundo*** (vide Figura 5 *ABSOLUTE DIFF*), a partir de uma imagem de referência (vide Figura 7). Como cada imagem é colorida, ou RGB (24 bits), a abordagem proposta pelo trabalho é a ***extração das três bandas*** (vide Figura 5 *EXTRACT 3 BANDS)* da imagem para processamento individual. Como cada banda da imagem constitui uma imagem em nível de cinza, é aplicado uma ***limiarização*** (vide Figura 5 *THRESHOLD(mmach))* produzindo imagens binárias. O próximo passo, é a utilização do operador ***gradiente morfológico*** (vide Figura 5 *MORPH. GRADIENT)* para extração das bordas das pessoas. Ao resultado da extração de bordas, se aplica o operador de ***fechamento de buracos*** (vide Figura 5 *CLOSE HOLES)*; e as pessoas são preenchidas. As três imagens binárias correspondentes ao tratamento de cada banda, são combinadas por operadores de ***união*** (vide Figura 5 *UNION(mmach))*. Partes brancas na imagem indicam onde as pessoas estão. Quando as pessoas se encontram perto uma das outras, dois ou mais pedestres são detectados como um único objeto. Sendo assim, é necessário um pós-processamento para corrigir este problema. Um operador de ***N-erosão*** (vide Figura 5 *N-EROSION)* e um operador de ***N-dilatação*** (vide Figura 5 *N-DILATION)* são aplicados em conjunto formando o que chamamos de ***filtro morfológico***. Este filtro é utilizado para remover pequenos objetos indesejáveis no processamento. Como penúltimo passo, um operador de ***N-abertura*** (vide Figura 5 *N-OPENING)* com um ***elemento estruturante*** (vide Figura 5 *3x3 Str. El.*) vertical 1x3 é aplicado para desconectar as pessoas. Finalmente, é efetuada uma ***rotulação*** (vide Figura 5 *LABELING(mmach))* onde cada pessoa presente na imagem recebe um nível de cinza; e através do bloco ***estatística*** (vide Figura 5 *STATISTICS)* é apresentado o número de pessoas presentes na imagem.

[1] Imagens RGB disponibilizam as informações das três bandas presentes (vermelha, verde e azul), cada uma delas equivalente a uma imagem monocromática. Embora se tenha maior quantidade de dados, evidentemente não se pode equiparar a três vezes a quantidade de informação, uma vez que boa parte pode ser redundante, em virtude da cena ser a mesma, obtida através de três filtros.

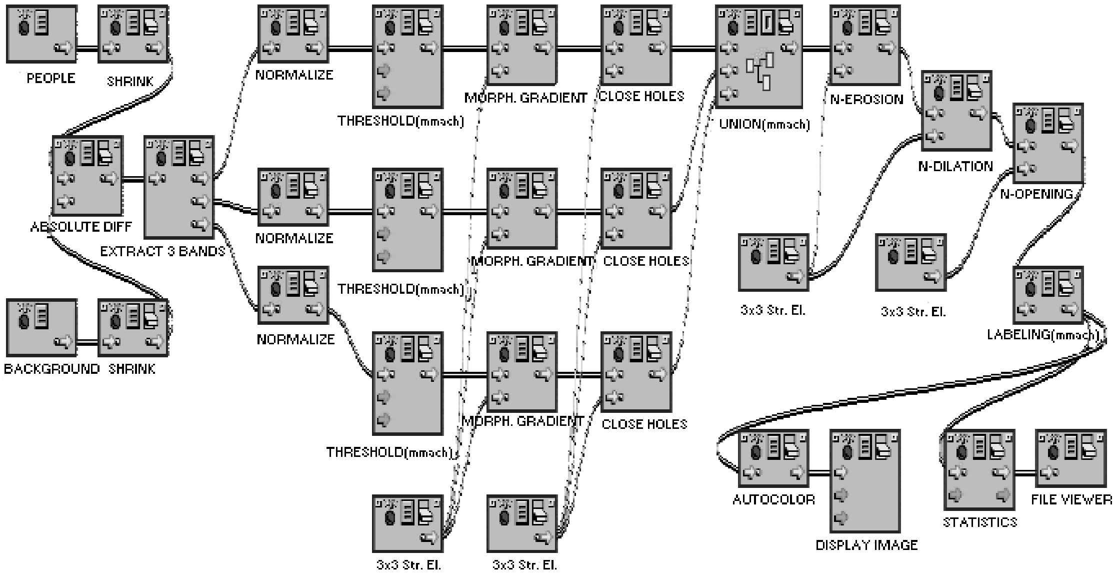
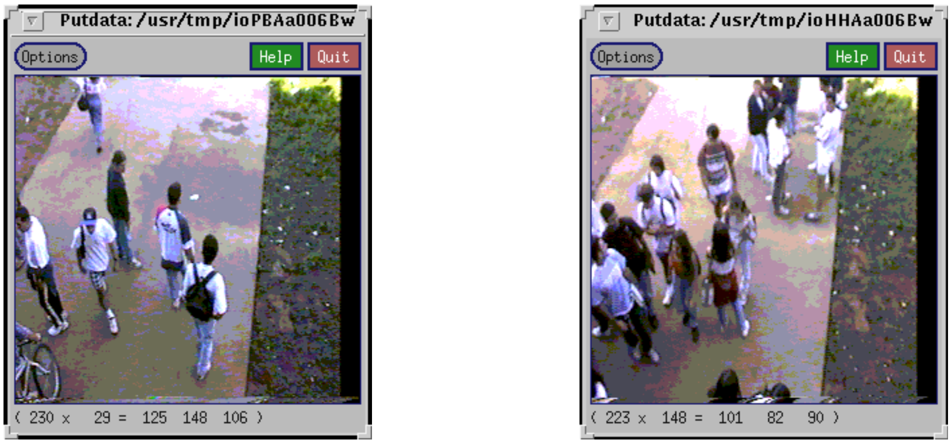


Figura 5. *Workspace* (ou programa visual) do Cantata, utilizado no processamento das imagens.



(a) (b)

Figura 6. Imagens usadas nos testes: (a) imagem com pouca quantidade de pessoas; (b) imagem com número mais elevado de pessoas presentes.

Figura 7. Imagem de referência, contendo somente o fundo.

A imagem da Figura 6a mostra pedestres se movimentando normalmente em um local no campus da Federal de São Carlos. Observa-se que a digitalização não é perfeita, pois existem manchas indesejáveis no fundo da imagem. A Figura 7 mostra a imagem de referência que foi utilizada para remover os objetos de fundo da imagem original. Nota-se no canto superior direito das duas imagens que as manchas devido aos raios solares não são exatamente iguais. Isto torna-se um problema pois a subtração não eliminará totalmente esta mancha. A Figura 8 mostra o resultado da subtração absoluta entre a primeira imagem teste (vide Figura 6a) e a imagem de referência Figura 7. Esta imagem é obtida após o bloco *ABSOLUTE DIFF* que pode ser visto na Figura 5. Observa-se que ocorre a eliminação da maioria dos objetos do fundo. Entretanto a imagem obtida é muito escura.

Figura 8. Diferença Absoluta

Figura 9. Limiarização

Pode-se verificar que após a subtração da imagem original pela imagem de referência ocorre a separação das três bandas da imagem resultante que é colorida. Como próximo passo é realizado a limiarização das imagens resultantes (banda vermelha, verde e azul). A Figura 9 mostra a limiarização da banda vermelha como exemplo; as outras bandas têm resultados similares. Nota-se a existência de alguns buracos nas pessoas nesta imagem. O passo seguinte é a extração de bordas das pessoas, a Figura 10 mostra o resultado da detecção das bordas das mesmas através da aplicação do operador gradiente morfológico.

Figura 10. Gradiente Morfológico

Figura 11. Fechamento de Buracos

A imagem da Figura 11 representa a operação de fechar buracos. Observa-se que as pessoas são quase totalmente preenchidas. O valor de limiarização utilizado foi oitenta, o número de erosões e dilatações foram duas e aberturas foram vinte e oito.

A imagem da Figura 12 mostra o resultado da união das três bandas da imagem original que passaram pelas mesmas operações. Com esta união aproveitamos as características das três bandas. Entretanto, pode-se observar na Figura 12 a existência de pequenos objetos indesejáveis. Para removê-los como mostra a Figura 13, foi utilizado um filtro morfológico que é composto por n-erosões e n-dilatações

Figura 12. União das três bandas

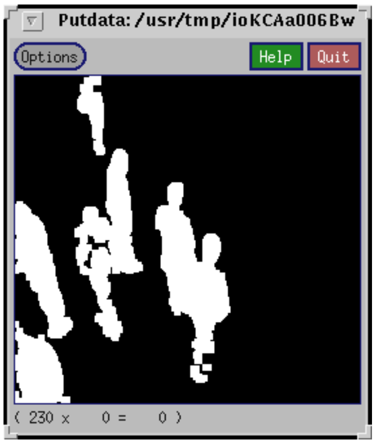


Figura 13. Remoção dos pequenos objetos

Figura 14. Rotulação

A última imagem apresentada é o resultado da rotulação das pessoas (vide Figura 14). Observa-se que as pessoas conectadas na Figura 13 foram desconectadas pela abertura e a cada pessoa foi atribuído um nível de cinza. Finalmente através do bloco estatística é apresentado o número de pessoas na imagem.

A segunda imagem teste (vide Figura 6b) apresenta uma densidade maior de pessoas. O maior problema é a sobreposição das mesmas devido a posição da câmera. Nota-se também o problema de digitalização da imagem.

O resultado da diferença absoluta continua sendo uma imagem muito escura (vide Figura 15), enquanto que o resultado da limiarização (vide Figura 16) mostra a região da imagem onde as pessoas se encontram mais alguns pequenos objetos. Nota-se na Figura 20 que estes pequenos objetos ou ruídos são removidos, porém o número de aberturas utilizadas para separar as pessoas sobrepostas não foram suficientes nesta densidade. Nota-se também na Figura 17, que o operador de gradiente morfológico extraiu relativamente bem o contorno das pessoas, e nas Figuras 18 e 19 vemos que os procedimentos de fechamentos de buracos e união das três bandas preenchem melhor as pessoas do que na Figura 16. A Figura 21 mostra o resultado da rotulação, onde cada pessoa recebe um nível de cinza correspondente.

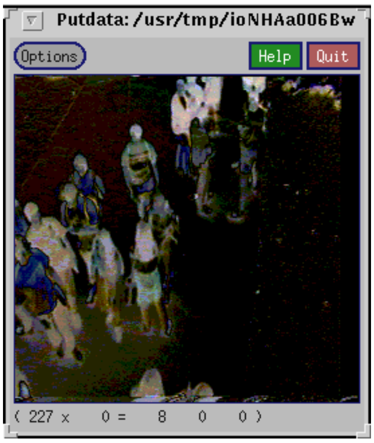


Figura 15. Diferença Absoluta

Figura 16. Limiarização

Figura 17. Gradiente Morfológico

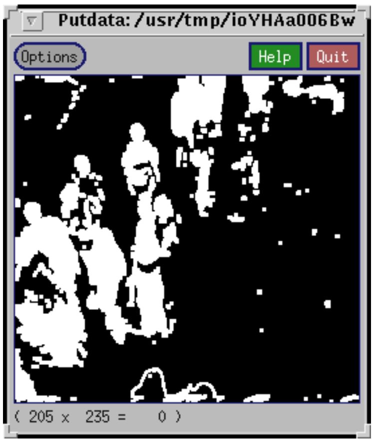


Figura 18. Fechamento de Buracos

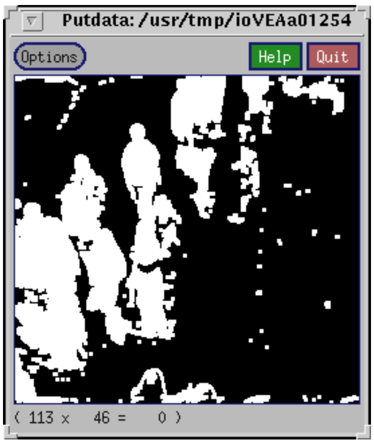


Figura 19. União das três bandas

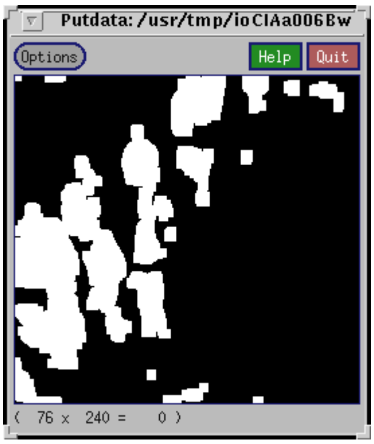


Figura 20. Remoção dos pequenos objetos

Pode-se perceber que a alta densidade de pessoas faz com que ocorra muita sobreposição. O nível utilizado para a limiarização foi oitenta, foram utilizadas três erosões e dilatações e quinze aberturas.

Pode-se observar que não foi perfeita a segmentação das pessoas devido a alta densidade de pedestres na imagem teste da Figura 6b. Os parâmetros envolvidos no processamento ainda são muito sensíveis a variação da densidade.

Figura 21. Rotulação

6 CONCLUSÃO

Os resultados obtidos mostram que esta abordagem tem produzido boas segmentações, contudo os parâmetros envolvidos ainda necessitam de ajustes para cada conjunto de imagens. A utilização de imagens coloridas tem produzido resultados melhores do que a utilização de suas versões monocromáticas, porque pode-se aproveitar informações das três bandas RGB. O tempo de processamento tem sido uma característica muita restritiva quando se está considerando controle em tempo real. Existe ainda uma grande sensibilidade quanto ao número de interações das aberturas para separar as pessoas sobrepostas.

Resultados obtidos neste trabalho foram apresentados no III Simpósio de Iniciação Científica em Ciências Matemáticas e da Computação na Universidade de São Paulo - Escola de Engenharia de São Carlos - Instituto de Ciências Matemáticas e no “Brazilian Workshop’96 on Mathematical Morphology” no IME - Instituto de Matemática e Estatística - Universidade de São Paulo.

7 PRÓXIMOS TRABALHOS

Para os próximos trabalhos, podem ser efetuadas análises com múltiplas seqüências de imagens. Esta nova análise poderia aumentar sensivelmente a qualidade dos resultados, principalmente em problemas de processamento em tempo real. Um estudo sobre a avaliação da direção do fluxo também poderá ser efetuado, melhorando a qualidade das informações obtidas.

Poderia ser intensamente explorado a utilização da cromaticidade dos dados, ou seja, explorar a viabilidade da utilização de operadores morfológicos que tratem imagens coloridas, as quais não se tem disponíveis atualmente.

8 AGRADECIMENTOS

Agradeço ao CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico pela bolsa de iniciação científica, ao Professor Jander Moreira pela orientação no desenvolvimento do trabalho e à EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária pela digitalização das imagens testes.

9 REFERÊNCIAS BIBLIOGRÁFICAS

[BANO94] BANON, G.J.F.; BARRERA, J. Bases da Morfologia Matemática Para a Análise de Imagens Binárias. IX Escola de Computação, Recife-UFPE-DI, 1994, 230p.

[GONZ92] GONZALES, R.C. & WOODS,R.E. “Digital Image Processing”. Addison Wesley Publishing Co., 2nd edition, 1992.

[POLU83] POLUS, A.; SCHOFER, J.L.; USHPZ, A. Pedestrian flow and level of service. *J. Transportation Eng.*, 109(1):46-56, Jan 1983

[SERR82] SERRA, J.P.F. “Image Analysis and Mathematical Morphology”. vol. 1, Academic Press - London, 610p. 1982.

[VELA93] VELASTIN, S.A.; YIN, J.H.; DAVIES, A.C.; VICENCIO-SILVA, M.A.; ALLSOP, R.E.; PENN, A. “Analysis of Crowd Movements and Densities in Built-Up Envvironments Using Image Processing”. **IEE Coll. on Image processing for Transport Aplications**, London, UK, 9 dec.: 8/1-8/6, 1993.

10 BIBLIOGRAFIA ADICIONAL

BARRERA, J.; BANON, G.J.F.; LOTUFO, R.A. “A Mathematical Morphology Toolbox for the KHOROS System”. **Image, Algebra and Morphological Image Processing V**, SPIE, San Diego - CA. 1994.

BARTOLINI, F.; CAPPELLINI, V.; MECOCCI, A. “Counting People Getting In and Out of a Bus by Real-Time Image-Sequence **Processing”. Image and Vision Computing**, Butterworth-Heinemann Ltd, **12** (1): 36-41, 1994.

JAIN, A.K. “Fundamentals of Digital Image Processing”. Prentice Hall, Englewood Cliffs, 1989.

JORDÁN, R. & LOTUFO, R. “Digital Image Processing with Khoros 2”, http://www.khoral.com/dipcourse http://dca.fee.unicamp.br/dipcourse (1994, 1995)

KHORAL RESEARCH, Inc., Khoros Manuals Vol. 1-9, 1994-1995.

LIM, J.S. Two- Dimensional Signal and Image Processing. Prentice-Hall, Englewood Clioffs - N.J., 1990.

SIBGRAPI 95 - VIII Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens., São Carlos-SP, 1995.