

Universidade Federal de Ouro Preto - UFOP
Instituto de Ciências Exatas e Biológicas - ICEB
Departamento de Computação - DECOM

CONTAGEM DE PESSOAS POR VÍDEO USANDO CÂMERAS EM POSIÇÃO ZENITAL

Aluno: Victor Hugo Cunha de Melo
Matricula: 08.1.4047

Orientador: David Menotti

Ouro Preto
11 de junho de 2011

Universidade Federal de Ouro Preto - UFOP
Instituto de Ciências Exatas e Biológicas - ICEB
Departamento de Computação - DECOM

CONTAGEM DE PESSOAS POR VÍDEO USANDO CÂMERAS EM POSIÇÃO ZENITAL

Relatório de atividades desenvolvidas apresentado ao curso de Bacharelado em Ciência da Computação, Universidade Federal de Ouro Preto, como requisito parcial para a conclusão da disciplina Monografia I (BCC390).

Aluno: Victor Hugo Cunha de Melo
Matricula: 08.1.4047

Orientador: David Menotti

Ouro Preto
11 de junho de 2011

Resumo

Detecção, rastreamento e contagem de pessoas são de grande utilidade para diversas aplicações comerciais, como monitoramento de espaços públicos, estádios de futebol, ou estações de ônibus. Neste projeto, propõe-se estudar e implementar métodos de contagem de pessoas por vídeo usando câmeras em posição zenital (rotacionada em 180 graus), que preservam a privacidade das pessoas. Ainda, propõe-se avaliar densamente os métodos estudados com vários vídeos adquiridos em ambientes distintos.

Palavras-chave: Reconhecimento de padrões. Contagem de pessoas. Câmera zenital.

Sumário

1	Introdução	1
2	Justificativa	2
3	Objetivos	3
3.1	Objetivo geral	3
3.2	Objetivos específicos	3
4	Metodologia	4
5	Desenvolvimento	5
5.1	O Segundo Método	5
5.1.1	Extração do Movimento	5
5.1.2	Contagem de Pessoas	5
5.2	Base de Dados	6
6	Trabalhos Futuros	6
7	Cronograma de atividades	6

Lista de Figuras

1	Resultados do passo-a-passo do método de <i>multiple lines</i>	6
---	--	---

Lista de Tabelas

1	Cronograma de Atividades.	7
---	-----------------------------------	---

1 Introdução

Detecção, rastreamento e contagem de pessoas são de grande utilidade para diversas aplicações comerciais, como monitoramento de espaços públicos, estádios de futebol, ou estações de ônibus. Possui grandes implicações em segurança, e permite coletar informações dos sistemas de forma que possam ser utilizados para identificar padrões em tráfego por horário, otimizar agendamento de trabalhos, monitorar a efetividade de eventos, *etc.*

Além de sensores de imagens, formas mecânicas e outras tecnologias de sensores são utilizadas para contagem de pessoas. Os sistemas que utilizam contadores mecânicos, como catracas, contam apenas uma pessoa por vez e podem obstruir a passagem, ocasionando congestionamentos se há muitos transeuntes. Devido ao seu projeto, está sujeita a subcontagens. Sistemas que utilizam raios infravermelhos ou sensores de calor não bloqueiam as portas, mas não apresentam precisão para identificar pessoas em um grupo. É notável a necessidade de sistemas mais informativos, por isso câmeras foram selecionadas como instrumento de detecção.

A segmentação do plano de fundo é o primeiro passo em diversas aplicações de visão computacional. Usualmente é obtido em sistemas de detecção humana calculando a diferença pixel-a-pixel entre o *frame* atual e a imagem do plano de fundo, seguido por uma limiarização automática. Se a precisão que essa abordagem oferece não é garantida, uma estratégia por blocos é preferível pois produz uma segmentação mais estável na presença de mudanças de luz e sombra.

Existem vários métodos propostos na literatura para a contagem de pessoas por vídeo [Bescos et al., 2003, Chien et al., 2004, Huang and Chow, 2003, Snidaro et al., 2005, Velipasalar et al., 2006]. No entanto, é difícil encontrar uma validação efetiva e extensa dos métodos propostos. Ainda, estes trabalhos não levam em conta a privacidade das pessoas sendo filmadas.

Neste projeto, propõe-se estudar e implementar métodos de contagem de pessoas por vídeo usando câmeras em posição zenital¹ [Antic et al., 2009, Barandiaran et al., 2008, Chen et al., 2008]. Ainda, propõe-se avaliar densamente os métodos estudados com vários vídeos sob diversas condições.

¹rotacionada azimutalmente em 180 graus

2 Justificativa

O projeto em questão é relevante tanto nas esferas social e ambiental, quanto na computacional.

O produto final que pode ser gerado com a conclusão deste projeto é um sistema de controle de acesso de pessoas as dependências do DECOM. Tal sistema, por meio da quantificação das pessoas nas dependências do DECOM, pode oferecer maior segurança aos alunos, professores e funcionários da universidade. Ainda, aumentará a visibilidade dentro da universidade tanto da pesquisa (PROPP e PPGCC) quanto do Departamento de Computação (DECOM).

Em caso de sucesso pleno, o projeto pode ainda ser encubado em uma empresa júnior para se tornar um produto construído dentro da própria UFOP a ser comercializado no Brasil.

Além disso, este projeto demonstra o potencial da área de processamento de imagem nas mais diversas aplicações, e seus resultados poderão servir de estímulo para que áreas, além da automação, passem a fazer uso de seus recursos.

Finalmente, conceitos da área de processamento de imagem, visão computacional e reconhecimento de padrões podem ser expandidos durante o estudo da literatura. O anseio maior deste projeto é obter um novo método para a solução desse problema particular: a contagem de pessoas por vídeo.

3 Objetivos

3.1 Objetivo geral

O objetivo geral deste projeto é pesquisar, caracterizar e implementar métodos para a contagem de pessoas por vídeo usando câmeras em posição zenital.

3.2 Objetivos específicos

Os objetivos específicos a serem atingidos são:

1. Fazer uma revisão da literatura sobre métodos de contagem de pessoas por vídeo usando câmeras em posição zenital.
2. Fazer uma revisão da literatura sobre processamento de imagem, visão computacional e reconhecimento de padrões, visando:
 - (a) Representação digital de imagens;
 - (b) Métodos de filtragem de imagem;
 - (c) Técnicas de identificação de objetos em imagens;
 - (d) Técnicas de remoção de *background*;
 - (e) Técnicas de rastreamento de objetos em vídeo;
3. Implementar os métodos de contagem de pessoas por vídeo;
4. Comparar e analisar os resultados obtidos pelos diferentes métodos implementados tendo como entrada vários vídeos obtidos nas dependências do DECOM em diversas épocas do dia e do ano;
5. Contribuir com a divulgação de técnicas de processamento gráfico / imagens e vídeo à solução de problemas de automação;
6. Produzir um artigo que possa ser publicado em um evento científico nacional e outro internacional e ainda outro que possa ser submetido a revista especializada;

4 Metodologia

As principais atividades previstas para esse projeto são:

1. Pesquisa de técnicas de processamento de imagem, reconhecimento de padrões e visão computacional e métodos para contagem de pessoas por vídeo;
2. Caracterização e classificação de cada método pesquisado;
3. Implementação dos métodos pesquisados;
4. Realização de testes em vídeos obtidos a partir das câmeras instaladas nas dependências do DECOM adquiridos em diversas épocas do dia e do ano;
5. Organização dos resultados obtidos;
6. Análise dos resultados;
7. Possível proposição e implementação de um novo método;
8. Preparação de artigos e pôsteres.

Considerando as atividades descritas, de forma geral, a metodologia prevista, para cada atividade discutida é a seguinte:

1. Inicialmente será feito um estudo das técnicas de processamento de imagem, reconhecimento de padrões e visão computacional, seguido de estudo de métodos para a contagem de pessoas por vídeo. Durante este processo estar-se-á classificando os métodos estudados;
2. Em seguida, os métodos estudados serão avaliados, implementados e testados;
3. Por fim, validar-se-á os métodos submetendo a eles vídeos contendo pessoas, visando avaliar a precisão dos métodos estudados. Espera-se também propor uma metodologia para avaliação automática dos métodos. Dessa forma, espera-se poder testar e validar a eficácia dos métodos estudados.

Observação: Para a realização dessas atividades, o projeto conta com o Laboratório de Processamento Digital de Imagem (LaPDI), localizado à sala COM20 do ICEB, com toda a infra-estrutura necessária ao desenvolvimento do projeto, no caso, microcomputadores e *softwares* “livres”, *scanner*, câmeras digitais e servidores com alta capacidade de armazenamento. Uma câmera necessária para aquisição dos vídeos já está instalada no corredor da secretaria do DECOM. Ainda, o laboratório conta com bibliografia especializada (livros internacionais) sobre processamento de imagem, reconhecimento de padrões e visão computacional.

5 Desenvolvimento

Anteriormente, já havíamos implementado o método proposto em [Antic et al., 2009]. Para comparar sua precisão com outros métodos de contagem de pessoas, optamos por implementar a solução proposta por [Barandiaran et al., 2008].

5.1 O Segundo Método

A idéia principal por trás desta solução consiste em definir uma área de interesse (*ROI*, *region of interest*) nas imagens onde analisamos os movimentos das pessoas. Inserimos linhas virtuais ortogonais a direção de movimento.

O algoritmo é dividido em três passos diferentes. Primeiro detectamos o movimento e extraímos a região onde as pessoas passaram. Após este passo, uma contagem cumulativa é efetuada pelas linhas virtuais. Finalmente, cada linha é examinada para a contagem de pessoas.

5.1.1 Extração do Movimento

Para extrair as informações de movimentação do plano de fundo, o método utiliza a diferença entre dois *frames* consecutivos. Utiliza-se um limiar arbitrário para filtrar os ruídos da imagem final.

5.1.2 Contagem de Pessoas

A segunda parte do algoritmo consiste na contagem realizada independentemente para cada linha pertencente a área de interesse das imagens. Cada linha é representada por uma função l , onde os eixos x e y correspondem a posição dentro da linha com o número acumulado de *pixels* da imagem.

$$l_x^t = l_x^{t-1} + D^t \quad \forall x \wedge 0 \leq x < czw \quad (1)$$

$$l_x^0 = 0 \quad \forall x \wedge 0 \leq x < czw \quad (2)$$

onde x é o ponto na linha; czw é a largura na área de interesse, que é igual a largura das linhas.

Os *pixels* são acumulados a cada vez que uma pessoa passa através de uma linha (função 1).

A figura 1 mostra estes passos do algoritmo. A coluna (a), exibe as imagens originais do vídeo. A coluna (b), ilustra a detecção de movimento de algumas pessoas por meio deste método. Inicialmente, não há linha nenhuma linha virtual pois ninguém cruzou a região de interesse. As linhas surgem quando uma pessoa cruza a região.

Não temos o resultado dos testes porque faltam detalhes sobre a parte final do algoritmo para concluirmos sua implementação. Estamos estudando suas referências para compreensão desta parte.

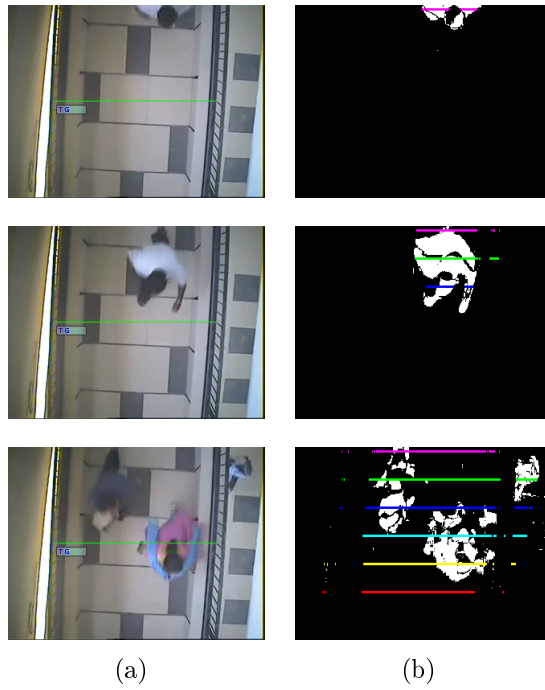


Figura 1: Resultados do passo-a-passo do método de *multiple lines*.

5.2 Base de Dados

Para a execução dos testes, gravamos cinco vídeos para compor nossa base de dados. A câmera, situada no corredor do Departamento de Computação, capturou vídeos de duração de uma hora, durante o dia e a noite. Os vídeos tem resolução de 640×480 , e utilizam o padrão JPEG.

6 Trabalhos Futuros

Planejamos como trabalhos futuros:

1. Concluir a implementação do método proposto no artigo [Barandiaran et al., 2008] e implementar outros métodos para avaliação.
2. Propor modificações nestes métodos estudados para aumentar a precisão.
3. Utilizar a nova base de vídeos para execução dos testes.
4. Estudar os métodos de rastreamento baseado em “*Natural Computing*”, como Filtro de Partículas e Colônia de Formigas para avaliar sua precisão.

7 Cronograma de atividades

Na Tabela 1, são apresentadas o planejamento das atividades que serão realizadas.

Atividades	Mar	Abr	Mai	Jun	Jul
Implementação	V	V	V	X	
Testes e Análises		V	V	X	
Redigir a Monografia			X	X	X
Apresentação do Trabalho					X

Tabela 1: Cronograma de Atividades.

Referências

- B. Antic, D. Letic, D. Culibrk, and V. Crnojevic. K-means based segmentation for real-time zenithal people counting. In *IEEE International Conference on Image Processing (ICIP)*, pages 2565–2568, 2009.
- Javier Barandiaran, Berta Murguia, and Fernando Boto. Real-time people counting using multiple lines. In *International Workshop on Image Analysis for Multimedia Interactive Services*, pages 159–162, 2008.
- J. Bescos, J. M. Menendez, and N. Garcia. DCT based segmentation applied to a scalable zenithal people counter. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 1005–1008, 2003.
- Chao-Ho Chen, Yin-Chan Chang, Tsong-Yi Chen, and Da-Jinn Wang. People counting system for getting in/out of a bus based on video processing. In *International Conference on Intelligent Systems Design and Applications (ISDA)*, pages 565–569, 2008.
- Shao-Yi Chien, Yu-Wen Huang, Bing-Yu Hsieh, Shyh-Yih Ma, and Liang-Gee Chen. Fast video segmentation algorithm with shadow cancellation, global motion compensation, and adaptive threshold techniques. *IEEE Transactions on Multimedia*, 6(5): 732–748, 2004.
- D. Huang and T. W. S. Chow. A people-counting system using a hybrid rbf neural network. *Neural Processing Letters*, 18:97–113, 2003.
- L. Snidaro, C. Micheloni, and C. Chiavedale. Video security for ambient intelligence. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 35(1):133–144, 2005.
- Senem Velipasalar, Ying li Tian, and Arun Hampapur. Automatic counting of interacting people by using a single uncalibrated camera. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1265–1268, 2006.