

# Rastreamento de múltiplos veículos em vias urbanas por processamento de vídeo

Gabriel I. S. Ruas, Ticiano A. C. Bragatto, Victor A. P. Benso, Marcus V. Lamar, Daniel Aldigueri, George L. Teixeira e Yaeko Yamashita

**Resumo**—Este trabalho apresenta uma nova metodologia baseada em processamento de vídeo para a contagem de veículos em cruzamentos e vias urbanas, visando mostrar a viabilidade das técnicas de visão computacional para a criação de um equipamento portátil capaz de realizar automaticamente contagem com discriminação de rotas de veículos e em tempo real. Técnicas de processamento de imagem, tais como subtração adaptativa de fundo, definição e rastreamento de características são usadas. Os resultados obtidos mostram a viabilidade do uso do sistema de monitoramento proposto em tempo real, alcançando taxas de acerto de 96% em vias simples e 72% em cruzamentos complexos.

**Palavras-Chave**—Contagem de veículos, visão computacional, cantos de Harris

**Abstract**—This work presents a new methodology based on video processing for recording and counting vehicles in intersections and urban roads. The aim of this work is to show the viability of computer vision techniques for the construction of a portable equipment able to perform urban vehicles flow counting and classification into routes in an automatic way and in real time. Image processing techniques, such as background subtraction, definition and tracking of object features are used. The obtained results show the viability of use of the proposed portable monitoring system in real-time, achieving correct rates in the vehicle counting of about 96% for simple roads and 72% at complex intersections.

**Keywords**—Vehicle counting, computer vision, Harris corners

## I. INTRODUÇÃO

O controle e gerência de tráfego em grandes e médias cidades são essenciais para garantir uma alta qualidade de serviço do sistema de transporte e para subsidiar uma avaliação da capacidade do sistema de atender a demanda. Este controle somente é possível através do monitoramento do atendimento da demanda, representado pelo uso do sistema pela população, isto é, pelo uso da oferta do sistema de transporte.

Existem diversas maneiras de se realizar esta tarefa, incluindo contagem manual, uso de laços indutivos e, mais recentemente, utilizando visão artificial. Utilizando visão computacional, pode-se realizar novas tarefas tais como detecção de colisão, classificação de veículos, estimação de velocidade, assim como rastreamento e contagem de veículos com ou sem discriminação de rotas.

Gabriel Ruas, Ticiano Bragatto, Daniel Aldigueri e George Teixeira, CEFTRU, Universidade de Brasília, Brasília, Brasil, E-mails: ruas@unb.br, bragatto@unb.br, dra@ceftru.unb.br, george@ceftru.unb.br.

Victor Benso, Dept. of Computer and Electrical Eng., Nagoya Institute of Technology, Nagoya, Japão, E-mail: vbenso@gmail.com.

Marcus Lamar, Dept. de Ciência da Computação, Universidade de Brasília, Brasília, Brasil, E-mail: lamar@unb.br.

Yaeko Yamashita, Dept. de Engenharia Civil, Universidade de Brasília, Brasília, Brasil, E-mail: yaeko@ceftru.unb.br.

Este trabalho se concentra nos problemas de rastreamento e contagem de veículos em rotas pré-definidas. Uma característica inovadora do trabalho ora proposto é a portabilidade do sistema.

## II. TRABALHOS ANTERIORES

Um sistema automático baseado em informações visuais deve ser capaz de detectar a presença de um veículo, seguir sua rota e contabiliza-lo em uma matriz Origem/Destino (O/D). Os principais desafios encontrados são: a separação dos objetos a serem rastreados da cena de fundo, oclusão de veículos por outros veículos, variações da luminosidade do ambiente devido às condições climáticas e a passagem do dia, e ruído devido a presença de outros objetos na cena.

Uma classe de abordagens é o modelamento tridimensional da cena [1], [2]. Nestes sistemas uma descrição precisa do movimento dos veículos pode ser encontrada, porém é necessário conhecer as coordenadas exatas das câmeras (geralmente 2 ou 3) e calibrar cuidadosamente o sistema para que seja possível analisar o espaço tridimensional. Isto dificulta o desenvolvimento de um sistema que possa ser usado em diferentes ambientes, uma vez que a portabilidade do sistema proposto pode leva-lo a cenários de difícil modelamento.

Muitos sistemas de visão monocular, que utilizam apenas uma câmera para rastrear objetos móveis em um sequência de vídeo têm sido propostos [3], [4], [5], [6], [7]. Estes sistemas podem ser divididos basicamente em duas categorias: rastreamento de contornos e rastreamento de características.

Sistemas que utilizam contornos para seguir objetos são amplamente descritos na literatura [8], [9], [10] e encontram aplicação em muitas áreas como rastreamento de olhos, lábios e pessoas. Para o caso de rastreamento de veículos com uma câmera fixa, os contornos podem ser usados eficientemente uma vez que eles podem ser estimados. Porém, com a câmera em uma posição arbitrária, modelar os contornos dos veículos é uma tarefa difícil. Portanto, sem um grande poder computacional, o uso de contornos não se apresenta como uma alternativa viável para a tarefa de rastreamento em sistemas portáteis.

Outro conjunto de métodos amplamente utilizado em aplicações de rastreamento visual de objetos tem como base a definição e rastreamento de características locais [11], [12], [13], [14]. Nesta classe de sistemas, a região que define um objeto não é rastreada diretamente no espaço da imagem, mas em um espaço transformado, no qual as características que o definem são robustas a deformações, mudanças do ponto de vista e variações de iluminação. O desenvolvimento de

técnicas eficientes de extração destas características ainda é um grande desafio para a comunidade científica. Normalmente, estas técnicas envolvem um processamento de alto custo computacional, o que exige o uso de processadores mais poderosos, aumentando o custo do projeto ou mesmo inviabilizando uma aplicação em tempo real com a tecnologia disponível. Um dos métodos mais populares e computacionalmente simples, para uso em aplicações de tempo real, é proposto por Harris [15] e é utilizado neste trabalho de maneira otimizada para caracterizar e rastrear um veículo.

### III. METODOLOGIA PROPOSTA

A proposta deste trabalho é a concepção e teste de um sistema portátil de contagem de veículos que possa ser facilmente instalado em diferentes posições em relação às vias monitoradas. A metodologia proposta por Cucchiara [14], que divide o sistema em módulos de processamento de imagem e de aplicação de regras, se mostrou adequada e foi adotada no desenvolvimento deste artigo. O sistema proposto, apresentado na figura 1, pode ser dividido nas seguintes etapas: Aquisição de imagem, subtração de fundo, segmentação e identificação de blobs, definição e rastreamento de características, e contagem dos veículos em uma matriz O/D.

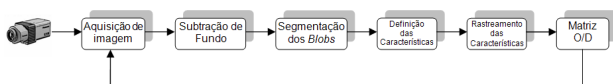


Fig. 1. Diagrama de blocos do sistema proposto

#### A. Aquisição de Imagens

No sistema proposto, imagens são adquiridas utilizando uma câmera de vídeo digital, cujo posicionamento é importante para o sucesso do monitoramento. A câmera deve ser posicionada de maneira a permitir que as imagens obtidas cubram toda a área de interesse da rua ou cruzamento, buscando minimizar a ocorrência de oclusões dos veículos e obstruções por outros objetos que possam fazer parte da cena. O posicionamento ideal é exatamente sobre o centro do cruzamento ou rua em uma vista superior da mesma, a uma altura suficiente para que todos os pontos de entrada e saída de interesse estejam visíveis [3]. Porém, este posicionamento não pode ser obtido de uma maneira rápida e simples na maioria dos cruzamentos. Portanto, um poste de 4,0 m de altura é utilizado como suporte para a câmera, permitindo uma boa visibilidade das vias, assim como dos pontos de entrada e saída a serem analisados.

#### B. Subtração de Fundo

A subtração de fundo é um método eficiente para a detecção de objetos em movimento em sequências de vídeo que tem ganhado a atenção de pesquisadores. Esta técnica envolve a estimação de uma imagem de referência a partir de uma sequência de imagens obtidas anteriormente, representando a característica estática da cena. As imagens capturadas são

então subtraídas deste fundo estimado e o resultado é comparado a um limiar, gerando uma máscara binária que destaca as regiões que correspondem às partes não-estacionárias da cena, isto é, possíveis objetos móveis. Em ambientes controlados sem variações de luminosidade, uma imagem real do fundo pode ser usada para este propósito, porém isto não é possível em aplicações de monitoramento automático, uma vez que o cenário muda constantemente, devido a, por exemplo, luzes piscando, árvores balançando, nuvens, e mudanças abruptas e graduais de luminosidade.

O método da mistura de Gaussianas para estimação de fundo, proposto originalmente por Stauffer [16], foi melhorado por KaewTraKulPong [17] e é usado neste trabalho devido à sua excelente robustez e adaptabilidade com a cena. Este método busca modelar os pixels de fundo por uma função densidade de probabilidade aproximada por uma soma ponderada de distribuições Gaussianas. A probabilidade que um pixel possua um valor  $x_N$  em um momento  $N$  pode ser modelada pela soma de  $K$  Gaussianas ponderadas,

$$p(x_N) = \sum_{k=1}^K w_k \eta(x_N; \mu_k, \Sigma_k) \quad (1)$$

onde  $w_k$  é o peso da  $k$ -ésima Gaussianas  $\eta(x; \mu_k, \Sigma_k)$ , definida por

$$\eta(x; \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \quad (2)$$

onde  $\mu_k$  é a média e  $\Sigma_k$  é a matriz de covariância da  $k$ -ésima Gaussianas.

Os pesos de cada Gaussianas representam a proporção de tempo que cada componente de cor está presente na cena. Desta maneira, o fundo é representado pelas Gaussianas que aparecem mais frequentemente na cena.

O método utilizado para adaptar cada Gaussianas (parâmetros  $w_k, \mu_k, \Sigma_k$ ) às variações do fundo da cena é uma variação do algoritmo de Expectation-Maximization (EM). O EM é um algoritmo iterativo que garante a convergência a um máximo local em um espaço de busca dado. Devido aos requisitos do modelamento do fundo, uma versão on-line do EM é usada, isto é, o modelo é ajustado progressivamente ao longo do tempo com a captura de novos quadros. Quando o fundo é modelado, regiões estáticas são representadas por grupos de Gaussianas mais concentradas, enquanto que regiões de movimento são representadas por Gaussianas mais dispersas. A atualização do modelo é realizada por um esquema de atualização seletiva, no qual cada novo pixel é comparado com os elementos do modelo existente, ordenados pelo peso. Se o valor do pixel estiver a uma distância ao centro menor ou igual a um limiar, calculado em função do desvio padrão da Gaussianas, a mesma é atualizada pelo algoritmo. Se o pixel não corresponde a nenhum dos componentes do modelo, a Gaussianas com a menor probabilidade é descartada e uma nova Gaussianas é criada, com a média igual ao valor do pixel, desvio padrão com alto valor e peso pequeno. A subtração é efetuada marcando todos os pixels que não foram classificados como

uma das cores do fundo como pertencendo à frente da cena. Esta abordagem permite que variações pequenas e rápidas das condições da filmagem sejam ignoradas. Porém, variações graduais, como a variação da iluminação ao longo do dia, são corretamente consideradas no modelo do fundo.

### C. Segmentação e Identificação de Blobs

O resultado obtido na subtração de fundo é uma imagem binária que representa as regiões de interesse, isto é, regiões que podem ser veículos em movimento. O objetivo do estágio de segmentação é separar e localizar estes objetos, além de filtrar objetos criados devido ao ruído ou objetos que não sejam de interesse.

O primeiro passo é remover regiões muito pequenas, criadas por ruído no sistema de aquisição, e destacar regiões que provavelmente representam objetos. Isto é feito aplicando-se filtros morfológicos de dilatação e erosão [18]. Este procedimento, além de eliminar falsos objetos, reforça as regiões conexas encontradas, eliminando lacunas que possam existir no interior dos objetos e evitando sua fragmentação. Uma vez que a máscara binária seja processada e o ruído eliminado, é necessário separar os objetos para o próximo estágio. Os objetos são localizados através do algoritmo de componentes conectados [18]. Este algoritmo percorre todos os pixels, localizando aqueles de mesmo valor e agrupando-os em regiões conexas, chamadas blobs neste artigo. As informações dos blobs obtidas após a análise de componentes conectados são a posição de sua centróide e as dimensões do retângulo que envolve essa região.

### D. Definição e rastreamento de Características

Para delimitar a região de processamento do estágio de rastreamento, as regiões de entrada e saída das vias no campo de visão da câmera são definidas manualmente durante a instalação e configuração do sistema no local a ser monitorado, conforme as figuras 2, 3, 4 e 5.

Inicialmente, os blobs identificados no estágio de segmentação são analisados. Uma nova entidade chamada *carro* é criada quando uma região correspondente a um blob se sobrepõe a uma entrada e o blob não cobre um *carro* já criado. Para cada *carro*, um conjunto de características é estimado e então rastreado. As características usadas neste artigo são os cantos de Harris [15]. O modelo de Harris é baseado na definição automática de pontos característicos (cantos e bordas) em uma região de interesse. As características extraídas pelo algoritmo são invariantes à iluminação, escalonamento e ruído na imagem gerando, portanto, um conjunto robusto. Os pontos característicos são definidos com base na função de autocorrelação da imagem. A função de autocorrelação mede as mudanças locais quando pequenos deslocamentos em diferentes direções são aplicados à imagem.

Considerando um deslocamento  $(\Delta x, \Delta y)$  e um ponto  $(x, y)$ , a função de autocorrelação pode ser definida como

$$c(x, y) = \sum_W (I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y))^2 \quad (3)$$

onde  $I$  corresponde à imagem e  $(x_i, y_i)$  são os pontos na janela Gaussiana ( $W$ ) centrada em  $(x, y)$ . A imagem deslocada  $I(x_i + \Delta x, y_i + \Delta y)$  pode ser aproximada truncando a série de Taylor nos termos de 1º grau,

$$I(x_i + \Delta x, y_i + \Delta y) \approx I(x_i, y_i) + \begin{bmatrix} I_x(x_i, y_i) & I_y(x_i, y_i) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (4)$$

onde  $I_x$  e  $I_y$  são as derivadas parciais nas direções  $x$  e  $y$  respectivamente.

Substituindo (4) em (3) obtemos

$$c(x, y) = \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} C(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (5)$$

onde  $C(x, y)$  corresponde à matriz que expressa a estrutura da vizinhança local, definida por

$$C(x, y) = \begin{bmatrix} \sum_W (I_x(x_i, y_i))^2 & \sum_W (I_x(x_i, y_i) I_y(x_i, y_i)) \\ \sum_W (I_x(x_i, y_i) I_y(x_i, y_i)) & \sum_W (I_y(x_i, y_i))^2 \end{bmatrix} \quad (6)$$

Estimando os autovalores  $\lambda_1, \lambda_2$  da matriz  $C(x, y)$ , uma descrição da região invariante à rotação é obtida. Com a análise dos valores de  $\lambda_1, \lambda_2$ , podemos determinar e caracterizar as regiões como aproximadamente uniformes ( $\lambda_1 \approx \lambda_2$  e valores pequenos), bordas verticais ( $\lambda_1 \gg \lambda_2$ ) ou horizontais ( $\lambda_1 \ll \lambda_2$ ), além da detecção de cantos ( $\lambda_1 \approx \lambda_2$  e valores altos). Desta maneira, podemos limitar a quantidade de características de interesse (cantos e bordas) para uma determinada região, ajustando um limiar ( $\xi_k$ ) para os autovalores.

No sistema proposto, o limiar de seleção dos autovalores ( $\xi_k$ ) para o procedimento de detecção de cantos de Harris é usado para adaptar a sensibilidade do detector de características às condições da via e dos veículos, de maneira a permitir um rastreamento consistente mesmo com variações das condições de iluminação ou cor aparente dos veículos.

Quando uma nova entidade *carro* é criada, o processo de identificação de características é feito na menor região retangular que contém o blob do objeto, de acordo com a metodologia descrita. Se o número de características encontradas for muito grande, apenas as 20 características mais significativas, isto é, com os maiores autovalores, são consideradas. Se a quantidade for muito pequena ( $< 3$ ), o valor do limiar  $\xi_k$  é reduzido e o procedimento de identificação de características é aplicado novamente. Este procedimento iterativo continua até que uma quantidade suficiente de características seja obtida, ou até que  $\xi_k$  alcance um valor mínimo pré-definido, quando o carro é descartado devido à alta complexidade de rastrear-lo, com um possível aumento no número de falsos negativos na contagem final. Felizmente, esta situação limite dificilmente é observada durante os monitoramentos de teste realizados.

Quando um *carro* já possui um grupo de características, devido ao rastreamento em quadros anteriores da sequência, o sistema aplica um algoritmo baseado no fluxo óptico estimado

entre o quadro atual e o quadro anterior, determinando o vetor de deslocamento de cada característica existente. O algoritmo aplicado neste artigo é o proposto por Lucas e Kanade [19] que, dentre as diversas propostas de algoritmos para estimação do fluxo ótico, possui a vantagem de possuir baixa complexidade computacional.

Uma vez que o deslocamento de todas as características entre os dois últimos quadros tenha sido obtido, as mudanças que ocorrem na entidade *carro* devido a este deslocamento são estimadas. Cada *carro* delimita uma região retangular do quadro, definida como a menor região que contém todas as características rastreadas. Considerando o deslocamento do veículo, existe a necessidade de estimar a sua posição no quadro seguinte, onde deve-se buscar pelas características.

Outra medida que deve ser estimada é a mudança de tamanho sofrida pelo *carro* entre os quadros devido à perspectiva. Para realizar estas estimativas, propõe-se o uso das características que apresentem maior consistência nos seus deslocamentos. A média e o desvio padrão dos deslocamentos sofridos pelas características pertencentes a um *carro* entre dois quadros consecutivos são estimados. As características que tenham se movido mais que um limiar, definido com base no desvio padrão das amostras de deslocamento, são descartadas, devido ao fato de elas violarem o princípio que um veículo é um objeto rígido; portanto, elas podem ser consideradas resultantes de erros de estimação do deslocamento. O deslocamento médio das características é usado para estimar a nova posição do *carro*. A região correspondente ao *carro* é estimada como sendo o retângulo que contém todas as características que foram consideradas consistentes. Com esta metodologia, evita-se a dispersão das características para regiões de fundo ou para outros veículos durante seu movimento, mantendo uma descrição compacta e consistente do veículo rastreado.

O rastreamento de todos os veículos (*carros*) presentes na cena continua até que eles atinjam uma das regiões de saída previamente definidas, quando sua rota é computada na matriz O/D. A matriz O/D apresenta o volume do fluxo de veículos, com as linhas representando as origens e as colunas representando os destinos.

#### IV. RESULTADOS

Como a proposta consiste de um sistema portátil de monitoramento, o mesmo foi implementado em um computador portátil com processador Intel CoreDuo 1.6 GHz, 1 GB RAM, 80 GB HD, sob o sistema operacional Linux. A programação foi feita na linguagem C, utilizando a biblioteca de visão computacional OpenCV. Os vídeos são capturados com uma câmera digital através da interface IEEE 1394, o que permite aquisição de imagens com resolução de 320x200 pixels, no espaço de cores YUV(4:2:2), a 30 quadros por segundo. O sistema proposto é capaz de processar o fluxo de vídeo de 30 fps (quadros por segundo) sem atrasos ou perda de quadros, portanto todas as etapas de processamento de imagem são executadas em menos de 33ms, utilizando o equipamento descrito, devido a baixa complexidade computacional das técnicas propostas.

Os resultados obtidos em quatro experimentos com diferentes números de vias, cruzamentos e faixas são apresentados nesta seção. Cada experimento de monitoramento possui uma duração de 15 minutos. O desempenho do sistema proposto é comparado com a contagem manual do fluxo de veículos realizada por um observador humano através de um vídeo gravado do mesmo ponto de vista da câmera. As matrizes O/D estimadas pelo sistema e as reais, obtidas através de contagem manual por um observador humano, são comparadas através do erro médio percentual de seus componentes, definido como

$$\bar{E} = \frac{100}{\|M_{OD}\|} \sum_{j=1}^M \sum_{i=1}^N |M_{OD}(i, j) - \tilde{M}_{OD}(i, j)| \quad (7)$$

onde  $M_{OD}$  e  $\tilde{M}_{OD}$  são as matrizes O/D real e estimada pelo sistema, respectivamente, contendo N entradas e M saídas.  $\|M_{OD}\|$  é a norma calculada pela soma dos elementos da matriz  $M_{OD}$ .

Para uma avaliação quantitativa da qualidade do sistema, além da comparação entre as matrizes O/D, a taxa de detecção correta (*Correct Detection Rate* - CDR) e taxa de alarme falso (*False Alarm Rate* - FAR), definidas como

$$CDR = \frac{TP}{TP + FN} \times 100 \quad (8)$$

e

$$FAR = \frac{FP}{TP + FP} \times 100 \quad (9)$$

onde TP (Positivos Verdadeiros) corresponde ao número de veículos rastreados e contados corretamente. FP (Falsos Positivos) representa o número de falsos alarmes positivos, isto é, casos em que o sistema rastreou um objeto através de uma rota válida, porém, o objeto não é um veículo, ou em casos nos quais o sistema separa um objeto sendo rastreado, considerando-o como dois ou mais objetos. FN é a quantidade de falsos negativos, isto é, veículos que passaram pelo campo de visão da câmera, mas não foram detectados pelo sistema. Um caso típico de falso negativo é quando ocorre oclusão entre veículos. Os resultados a seguir são apresentados em ordem crescente de complexidade da cena.

Inicialmente, o sistema proposto é aplicado a um caso simples, composto de duas vias com duas faixas cada. A câmera é colocada em uma posição próxima da ideal, isto é, uma visão superior no centro das duas vias, a 6 metros de altura, conforme a Figura 2.

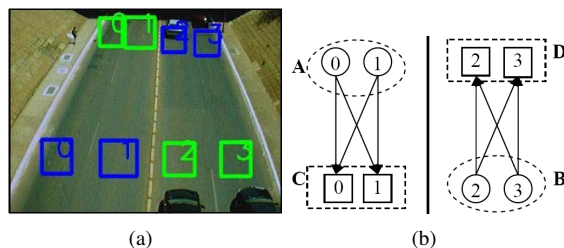


Fig. 2. Experimento 1, (a) cena real, (b) esquemático de rotas possíveis

A Figura 2(a) apresenta uma amostra da cena capturada pela câmera e a Figura 2(b) apresenta um esquemático sobre

as possíveis rotas dos veículos na cena. Círculos indicam os pontos de entrada e retângulos indicam os pontos de saída nas vias definidas durante a inicialização do sistema. Para analisar o fluxo de veículos na Via 1, as rotas  $0 \rightarrow 1$  e  $1 \rightarrow 0$  foram consideradas junto com as rotas básicas  $0 \rightarrow 0$  e  $1 \rightarrow 1$ , criando a rota  $A \rightarrow C$ , devido ao fato que o veículo realiza o mesmo movimento, apenas mudando de faixa. O mesmo é feito na Via 2, criando a rota  $B \rightarrow D$ . Desta maneira, o fluxo foi calculado considerando apenas duas entradas e duas saídas, e apenas duas rotas válidas, uma vez que é proibido pelas leis de trânsito fazer as rotas  $A \rightarrow D$  e  $B \rightarrow C$ .

A matriz O/D estimada pelo sistema possui duas origens (A e B) e dois destinos (C e D), sendo comparada à matriz obtida por contagem visual humana.

O erro médio percentual da matriz O/D estimado pelo sistema é de 4,04%, resultando em uma taxa média de acertos de 95,96%.

TABELA I  
RESULTADOS DO EXPERIMENTO 1

Rotas	TP	FN	FP	CDR	FAR
$A \rightarrow C$	119	8	0	93,7%	0,0%
$B \rightarrow D$	213	6	3	97,3%	1,4%
Total	332	14	3	96,0%	0,9%

A Tabela I apresenta o número de veículos identificados corretamente (TP), o número de falsos negativos (FN), o número de falsos positivos (FP), taxa de detecção correta (CDR) e a taxa de alarmes falsos (FAR). Falsos positivos ocorrem devido a erros na definição da entidade *carro* a partir do blob, o que fez que alguns veículos fossem rastreados como dois. Neste experimento, devido ao posicionamento próximo do ideal da câmera, não ocorrem oclusões entre os veículos. Falsos negativos ocorreram, principalmente, devido a movimentações súbitas da câmera, durante as quais o sistema não rastreou 14 veículos.

No segundo caso de estudo, o sistema de monitoramento foi empregado em uma via simples de mão única com duas faixas. A câmera está posicionada a 4,0 metros de altura, num ponto de vista perpendicular à borda da via, conforme apresentado na figura 3.

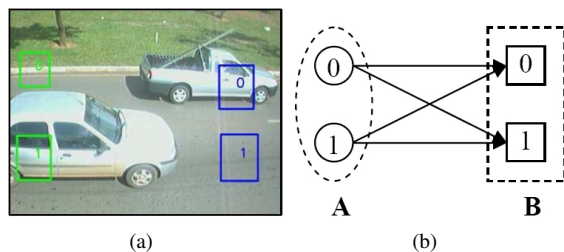


Fig. 3. Experimento 2, (a) cena real, (b) esquemático de rotas possíveis

Neste experimento, há a influência do sol, causando reflexões nas janelas dos veículos e saturando o sensor CCD da câmera. Outra dificuldade é a pequena distância entre a câmera e a via, o que faz com que veículos grandes como caminhões e ônibus ocupem toda a área de captura, dificultando o processo de identificação de características e aumentando a possibilidade de oclusão total e parcial de outros

veículos. Neste caso, ambas as faixas foram analisadas como possuindo apenas um ponto de entrada e um ponto de saída para estimar o fluxo de veículos. Portanto, a matriz O/D possui apenas um elemento correspondente ao fluxo  $A \rightarrow B$ . O erro médio percentual é de 14,8% na estimativa da matriz O/D. A tabela II apresenta resultados detalhados do experimento 2.

TABELA II  
RESULTADOS DO EXPERIMENTO 2

Rotas	TP	FN	FP	CDR	FAR
$A \rightarrow B$	188	8	37	95,9%	16,4%

Devido à pequena distância entre a câmera e a via, veículos grandes são, na maior parte dos casos, segmentados em duas ou mais partes, levando a um aumento significativo de falsos positivos. Estes mesmo veículos são os principais responsáveis pelos falsos negativos, devido à grande quantidade de oclusão das vias quando estes passam, impedindo a visão de carros menores que passem atrás deles.

Para avaliar a performance em um caso estruturalmente mais complexo, uma junção de duas vias de mão dupla foi monitorada, tal como na figura 4.

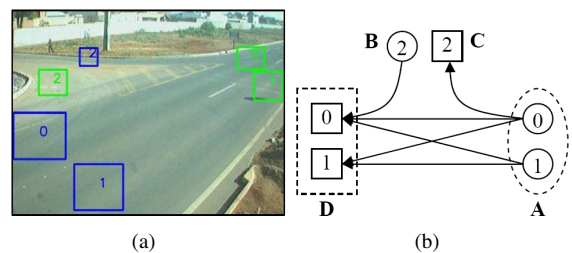


Fig. 4. Experimento 3, (a) cena real, (b) esquemático de rotas possíveis

Neste experimento, a ocorrência de oclusões foi notada frequentemente. O erro médio percentual obtido é de 8,77%.

TABELA III  
RESULTADOS DO EXPERIMENTO 3

Rotas	TP	FN	FP	CDR	FAR
$A \rightarrow C$	1	0	1	100,0%	50,0%
$A \rightarrow D$	46	10	3	82,1%	6,1%
$B \rightarrow D$	0	0	11	100,0%	100,0%
Total	47	10	15	82,5%	24,2%

O grande número de oclusões ocorreu devido a limitações no posicionamento da câmera e à presença de uma grande quantidade de ônibus. Mesmo assim, os resultados obtidos são bastante promissores, conforme o apresentado na tabela III, obtendo uma taxa de detecção correta de 82,5%.

No experimento 4, foi realizado o monitoramento das saídas e entradas de uma via de mão única com 3 faixas, como pode ser visto na Figura 5.

As rotas monitoradas resultam em uma matriz O/D composta de 2 entradas e 3 saídas. O erro médio percentual obtido foi de 24,0%.

Neste experimento, notou-se a robustez do sistema quando a câmera se encontra a uma distância suficiente do cruzamento e as condições ambientais (o sol e as sombras) não são críticos. A rota  $B \rightarrow D$  possui uma alta taxa de Falsos Negativos devido ao fato que os veículos têm que aguardar para entrar na



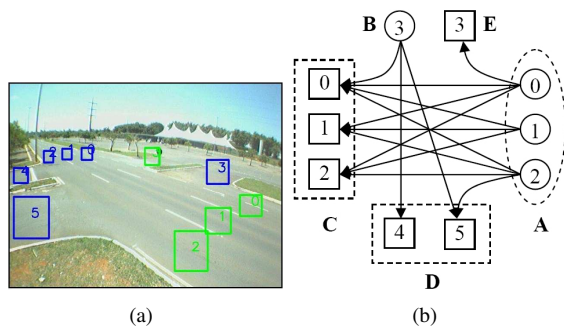


Fig. 5. Experimento 4, (a) cena real, (b) esquemático de rotas possíveis

TABELA IV  
RESULTADOS DO EXPERIMENTO 4

Rotas	TP	FN	FP	CDR	FAR
A → C	29	3	1	90.6%	3.3%
A → D	1	0	0	100.0%	0.0%
A → E	2	2	0	50.0%	0.0%
B → C	0	2	1	0.0%	100.0%
B → D	4	7	0	36.4%	0.0%
Total	36	14	2	72.0%	5.3%

pista. Caso este tempo de espera seja muito longo, o algoritmo de adaptação do fundo modela o veículo como sendo parte do fundo, causando a imprecisão. Neste caso, o sistema proposto alcançou uma taxa de acerto de 72,0% com uma taxa de alarmes falsos baixa, de 5,3%, conforme a tabela IV.

### V. CONCLUSÕES

Este artigo propôs o desenvolvimento de um sistema visual automático de determinação do fluxo de veículos em vias urbanas, tendo como principal diferencial a portabilidade, conseguida através do uso de um sistema adaptativo e robusto.

O sistema proposto é baseado na análise visual de um cruzamento pelo uso de um computador portátil e de uma câmera digital de baixo custo. Problemas inerentes à portabilidade do sistema, como variações da cena, das condições de iluminação, mudanças no formato do veículo dependendo da posição da câmera, e a presença de ruído e regiões de movimento indesejadas foram analisados. Soluções baseadas em técnicas de subtração adaptativa de fundo, definição e rastreamento de características foram aplicadas.

O protótipo do sistema foi desenvolvido e utilizado em situações reais de tráfego, mostrando ser eficiente e robusto, alcançando taxas de acerto de 96% quando comparados aos resultados obtidos por análise manual humana, em vias simples. Em cruzamento complexos, devido ao aumento da complexidade da cena, um desempenho de 72% de acerto foi obtido. A contagem manual foi realizada em um vídeo gravado, do mesmo ponto de vista que o sistema utiliza, ou seja, corresponde a quantidade real de veículos que o sistema automático deveria ser capaz de obter. Porém, o objetivo do sistema proposto é a realização da contagem em tempo real, situação na qual a contagem manual é realizada tipicamente por uma equipe de pessoas posicionadas ao lado da via, estando portanto esta contagem sujeita aos erros humanos devido a fadiga, faltas e distrações inerentes a trabalhos repetitivos. Neste trabalho não foi efetuada ainda esta comparação.

Para trabalhos futuros, este protótipo será melhorado para torná-lo mais robusto a variações de iluminação durante o dia, permitindo um monitoramento contínuo do fluxo de veículos por longos períodos de tempo, medidas que são impossíveis de se obter com as técnicas de contagem manual em uso atualmente. Com esta informação, gerentes urbanos e engenheiros podem desenvolver técnicas mais eficientes de gerenciamento de recursos e planejamento do transporte urbano, melhorando a qualidade de vida da população.

### REFERÊNCIAS

- [1] T. J. Ellis and M. K. Teal, "Occlusion prediction based on target dynamics and spatial reasoning," *Image Processing and Its Applications, 1997., Sixth Int. Conf. on*, vol. 1, pp. 219–223, 1997.
- [2] Z. Kim and J. Malik, "Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking," in *ICCV '03: Proc. of the Ninth IEEE Int. Conf. on Computer Vision*. Washington, DC, USA: IEEE Computer Society, 2003, p. 524.
- [3] S.-M. Lee, H. Baik, and J. H. Park, "Visual traffic movement counts at intersection and origin-destination (o-d) trip table estimation," in *Intelligent Transportation Systems Conf., 2007. ITSC 2007. IEEE, 2007*, pp. 1108–1113.
- [4] H. Unno, K. Ojima, K. Hayashibe, and H. Saji, "Vehicle motion tracking using symmetry of vehicle and background subtraction," in *Intelligent Vehicles Symposium, 2007 IEEE, 2007*, pp. 1127–1131.
- [5] Y. Yang, C. Yang, and X. Wang, "A detection and tracing algorithm of moving vehicles," in *Pervasive Computing and Applications, 2007. ICPCA 2007. 2nd Int. Conf. on*, 2007, pp. 447–450.
- [6] Z.-f. Liu and Z. You, "A real-time vision-based vehicle tracking and traffic surveillance," in *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007. SNPD 2007. Eighth ACIS Int. Conf. on*, vol. 1, 2007, pp. 174–179.
- [7] J.-W. Hsieh, S.-H. Yu, Y.-S. Chen, and W.-F. Hu, "Automatic traffic surveillance system for vehicle tracking and classification," *Intelligent Transportation Systems, IEEE Trans. on*, vol. 7, no. 2, pp. 175–187, 2006.
- [8] V. Gemignani, S. Proveddi, M. Demi, M. Paterni, and A. Benassi, "A dsp-based real time contour tracking system," in *Image Analysis and Processing, 1999. Proc. Int. Conf. on*, 1999, pp. 630–635.
- [9] A. Techmer, "Contour-based motion estimation and object tracking for real-time applications," in *Image Processing, 2001. Proc. 2001 Int. Conf. on*, 2001, pp. 648–651.
- [10] Q. Wang, H. Ai, and G. Xu, "A probabilistic dynamic contour model for accurate and robust lip tracking," in *Multimodal Interfaces, 2002. Proc. Fourth IEEE Int. Conf. on*, 2002, pp. 281–286.
- [11] H. S. Richardson and S. D. Blostein, "A sequential detection framework for feature tracking within computational constraints," in *Computer Vision and Pattern Recognition, 1992. Proc. CVPR '92., 1992 IEEE Computer Society Conf. on*, June 1992, pp. 861–864.
- [12] X. Xie, R. Sudhakar, and H. Zhuang, "Real-time eye feature tracking from a video image sequence using kalman filter," *Systems, Man and Cybernetics, IEEE Trans. on*, vol. 25, no. 12, pp. 1568–1577, 1995.
- [13] T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto, "Making good features track better," in *Computer Vision and Pattern Recognition, 1998. Proc. 1998 IEEE Computer Society Conf. on*, 1998, pp. 178–183.
- [14] R. Cucchiara and M. Piccardi, "Vehicle detection under day and night illumination," in *Proc. of the Third Int. Symposia on Intelligent Industrial Automation and Soft Computing, 1999*. [Online]. Available: citeseer.ist.psu.edu/cucchiara99vehicle.html
- [15] C. Harris and M. J. Stephens, "A combined corner and edge detection," in *Proc. of The Fourth Alvey Vision Conf.*, 1988, pp. 147–151.
- [16] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conf. on.*, vol. 2, 1999, p. 252 Vol. 2.
- [17] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," in *2nd European Workshop on Advanced Video Based Surveillance Systems*. Kluwer Academic Publishers, September 2001.
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1992.
- [19] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. of the 7th Int. Joint Conf. on Artificial Intelligence (IJCAI '81)*, April 1981, pp. 674–679.