

Deteccção e rastreamento de pedestres em travessias

David R. P. Gibson, Bo Ling, Michael Zeifman, Shaoqiang Dong, and Uma Venkataraman *

Publicado no Public Road Magazine, edição de Março/Abril de 2006

Um novo sistema de deteção, fundamentado no princípio da estereovisão, promete melhorar a segurança dos pedestres, no futuro.

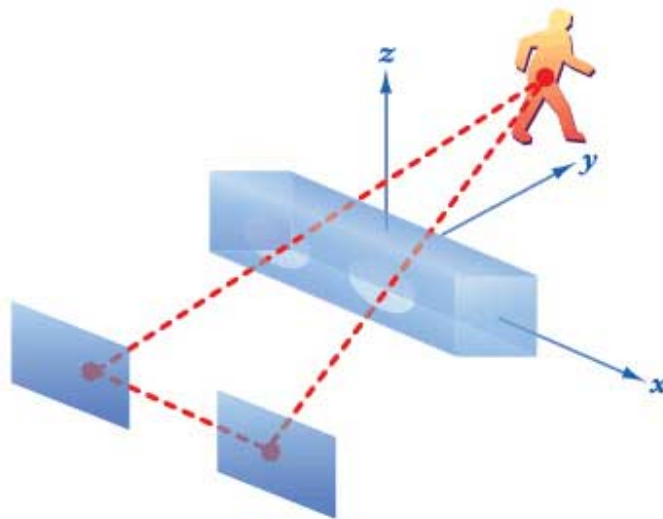


Nova pesquisa da FHWA, com o propósito de rastrear pedestres atravessando a via, pode vir a reduzir o número de atropelamentos. Foto: AAA Foundation for Traffic Safety

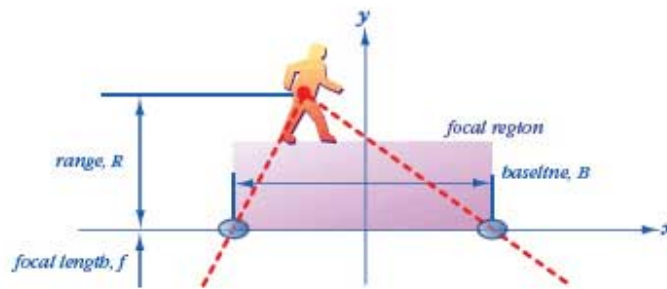
Todo ano, aproximadamente 5000 pedestres morrem em acidentes de trânsito nos Estados Unidos. A fim de melhorar sua segurança, os pesquisadores do FHWA estão se valendo das mais recentes tecnologias para detectar e rastrear os pedestres em suas travessias. O rastreamento computadorizado de pedestres oferece a possibilidade de desenvolver e instalar instrumentos de bordo nos veículos, que podem ajudar os condutores a perceber e evitar pessoas atravessando o leito viário. Esta técnica também pode ser aplicada ao dimensionamento automático do tempo de vermelho intermitente e vermelho de limpeza ao fim do estágio de pedestres.

Uma equipe de cinco pessoas, composta por pesquisadores do FHWA e por fabricantes de equipamentos, instalaram numa interseção duas câmeras de vídeo, controladas por um PC, que passaram a registrar a movimentação das pessoas. A parte mais complexa e importante da pesquisa consistiu em conseguir identificar cada uma das pessoas, distinguindo-a das demais e do ambiente do entorno. O sistema também foi capaz de calcular a velocidade dos pedestres e sua localização à medida que prosseguiam na travessia.

"Nós precisamos descobrir modos de conseguir reconhecer se há pedestres numa interseção e se necessitam ser atendidos", diz Tom Dodds, P. E., engenheiro especializado em pedestres e bicicletas do Departamento de Trânsito da Carolina do Sul (SCDOT). "Eu percebi, aqui na Carolina do Sul, que alguns pedestres não utilizam a boteira dos semáforos. Em condições desfavoráveis de visibilidade, como por exemplo à noite e sob chuva, é muito problemático, para o motorista, localizar esses transeuntes."



O diagrama mostra a configuração básica da estereo camera considerando as três dimensões, x , y e z . A caixa retangular representa a estereocamera equipada com duas lentes. Os outros dois retângulos representam os planos focais de cada lente, dentro da câmera. O plano que contém o triângulo tracejado é chamado de plano epipolar. Fonte: Migma Systems, Inc.



A figura traz uma visão 2-D da estrutura da estereo câmera. O segmento de reta localizado entre os focos direito e esquerdo é chamada de segmento de base, B (baseline). Geralmente, B é muito pequeno. Por exemplo, essa linha mede somente 12 centímetros no sistema montado pela pesquisa do FHWA. A distância entre o plano da imagem e o foco é denominada distância focal, f (focal length). A área entre os eixos focais direito e esquerdo é a região focal (focal region). O parâmetro de maior interesse é a distância entre o objeto e o segmento de base, R (Range). Fonte: Migma Systems, Inc.

A necessidade da detecção passiva (sem utilização de botoeira)

William C. Kloos, P.E., gerente do departamento de sinalização semafórica de Portland, OR, expõe sua opinião: "A detecção acurada e confiável de pedestres e condutores de bicicletas é a chave para incrementar o fator segurança na operação semafórica". A seguir, explica o porquê:

"Em muitos locais de Portland, os pedestres chegam a apertar a botoeira, mas se ocorrer uma brecha suficiente entre os veículos, irão atravessar mesmo no seu vermelho. Logo depois, tem início o estágio específico de pedestres, mas já não há mais ninguém para atravessar o que leva os motoristas a questionar porque foram obrigados a parar desnecessariamente. Se pudéssemos contar com uma detecção confiável de pedestres, poderíamos programar o cancelamento da demanda da botoeira nestes casos."

"Em muitos locais, os pedestres não compreendem que precisam apertar a botoeira para que apareça seu estágio. Com a detecção passiva de pedestres, não haveria mais necessidade de botoeiras. Nós, atualmente, utilizamos a detecção passiva para estender o intervalo de limpeza dos pedestres. Este método nos permite programar um intervalo no seu limite mínimo e estendê-lo para pessoas mais lentas."

"A detecção passiva pode também aprimorar os serviços que disponibilizamos para pessoas com necessidades especiais. Em muitas interseções existentes, as botoeiras estão bastante afastadas do meio-fio. Esta localização é ruim para pessoas com cadeiras de rodas ou com deficiência visual. O modo de transporte por bicicleta está,

cada vez mais, ganhando vulto. Com uma detecção eficiente de bicicletas, podemos otimizar a programação dos semáforos de modo a atendê-las especificamente.”

Recentemente, considerando os possíveis benefícios para pessoas com deficiência visual, o departamento United States Access Board disponibilizou sua última versão do Draft Guidelines for Accessible Public Rights-of-Way, que requisita a implantação de semáforos dedicados a essas pessoas em todas as novas implantações.

Desafios na detecção computadorizada de pedestres

A detecção de pedestres por vídeodetecção em ambientes a céu aberto constitui um complexo desafio. Embora tal processo possa parecer simples quando se pensa na visão do olho humano, a tarefa é extremamente complexa para um computador. Na visão por processamento de imagens, a imagem captada é dividida em elementos (pixels). Cada pixel traz consigo informações sobre cor e iluminação (denominada neste contexto como intensidade). A fim de conseguir identificar corretamente uma certa parte da imagem, o computador necessita recorrer a critérios de seleção de reconhecimento, como, por exemplo, a aparência do objeto. O processo é dificultado por vários fatores: pessoas utilizam roupas com cores que podem vir a se confundir com o entorno; podem usar chapéus ou carregar carrinhos de mão; podem estar paradas ou andando e, além disso, mudar de direção repentinamente. Para dificultar ainda mais a questão, o próprio entorno está permanentemente se modificando, tanto com a presença de elementos estáticos, como edifícios, vias, a própria sinalização de trânsito e os carros estacionados como pelos elementos móveis como, por exemplo, veículos em movimento, bicicletas e pedestres.

A abordagem tradicional da detecção computadorizada de pedestres se apóia na utilização de uma câmera comum. Atualmente, os algoritmos de processamento seguem uma das três seguintes metodologias: comparação com padrões pré-determinados, subtração do entorno e detecção do movimento.

Intelligent Vehicle Initiative

O trabalho do FHWA no tema “Rastreamento de pedestres” é um dos mais recentes estudos do Intelligent Vehicle Initiative (IVI), um programa de pesquisa e desenvolvimento do U.S. Department of Transportation's (USDOT) com ênfase em sistemas de segurança veicular e sistemas de informação ao motorista. IVI aglutina o esforço de vários órgãos: FHWA, National Highway Traffic Safety Administration e Federal Transit Administration. O programa USDOT's Intelligent Transportation Systems (ITS) Joint Program Office viabiliza um orçamento unificado para os órgãos relacionados. Desta forma, é possível promover a sinergia entre as pesquisas além de economizar os recursos necessários. O papel do IVI é o de pesquisar e desenvolver conceitos integrados que impliquem em maior segurança

viária e que, por não possuírem apelo comercial imediato para o mercado, não atraem os investimentos das empresas.

Os projetos do IVI têm se preocupado em investigar os fatores que envolvem o comportamento humano tais como a aceitação do usuário, o desenvolvimento de sistemas de informação ao condutor além de sistemas automáticos para evitar colisões e sistemas para vias automatizadas. Um crescente número de aplicações do IVI está se preocupando com o controle do trânsito baseado em detecção de pedestres, monitoração do fator segurança e análise do fluxo de pedestres.

Na abordagem que utiliza a comparação com padrões pré-determinados, é criada uma biblioteca contendo os possíveis padrões visuais de pedestres a fim de permitir o confronto entre uma parte da imagem do vídeo que está sendo analisada e cada um dos padrões armazenados. Quando é encontrada alguma semelhança, a parte da imagem correspondente é classificada como uma imagem de pedestre. Os pesquisadores do FHWA não consideram esta abordagem muito eficiente porque requer a criação de uma biblioteca com um enorme número de padrões visuais possíveis.

A meta da subtração do entorno é conseguir extrair a figura do pedestre do entorno em que está mergulhado. A chave, neste caso, é o conhecimento prévio da configuração do entorno do local. Se este é fixo, o pesquisador tem condições de, facilmente, "subtrair" o pedestre de imagens (frames) subsequentes. Quando isso não ocorre, é necessário recorrer a uma análise estatística de várias configurações possíveis. O maior defeito desta abordagem é a ausência de generalização, o que implica em que o modelo não pode ser validado se a configuração do entorno for relativamente nova.

Na abordagem da detecção do movimento, o pesquisador assume que os pedestres estão se movendo. Desta forma, todos os objetos em movimento são tratados como pedestres em potencial. Embora seja relativamente fácil detectar objetos em movimento dentro de imagens subsequentes (frames), os pesquisadores descobriram que não é tão fácil assim distinguir pedestres em movimento de outros objetos em movimento, como, por exemplo, veículos.

Tais abordagens tradicionais podem ser insuficientes para enfrentar as condições do mundo real, com sua enorme variedade de tipos de pessoas e de entornos não-estáticos. Recentemente, foi proposta uma nova abordagem, fundamentada na estereovisão tri-dimensional (3D). Esta técnica permite que sejam estimadas as distâncias entre elementos escolhidos, permitindo, assim, a identificação e o rastreamento de pedestres. Esta abordagem utiliza duas câmeras de vídeo que monitoram a mesma região do cruzamento desde dois pontos diferentes. A diferença entre os dois pontos de vista ocasiona um deslocamento relativo dos elementos observados nas imagens estereo. Explicando de outra forma, um mesmo ponto do cruzamento aparecerá levemente deslocado nas imagens tomadas por duas câmeras

próximas entre si. Tal deslocamento é quantificado por um parâmetro denominado "incoincidência", que constitui a base para o cálculo da distância entre os objetos e a câmera. Por exemplo, quanto mais afastado estiver um elemento das câmeras, menor será o valor da incoincidência. Normalmente, quando uma câmera monocular captura uma imagem 3-D, esta informação da distância é perdida. Com a estereovisão, entretanto, é possível deduzir que a distância até a parte superior de um elemento vertical, como é o caso do pedestre, é menor do que aquela até o chão atrás do elemento ou até os objetos mais distantes em volta. Este é o princípio sobre o qual se baseia a detecção de pedestres através da estereovisão.

Ainda que o recurso da estereovisão constitua um importante passo no sentido de descobrir objetos verticais no meio de todo o "ruído" do entorno, vários desafios ainda precisam ser superados até que se possa aplicar esta metodologia à detecção e rastreamento de pedestres, com a confiabilidade necessária.

Por que usar a incoincidência estereo?

A estereoscopia é um processo que permite determinar a distância entre um objeto e seu observador, ou em outros termos, que propicia a noção de profundidade. O termo tem origem em dois radicais gregos: stereo que significa solidez e oipsis, que quer dizer visão. A estereovisão computadorizada requer a comparação das informações 3-D de duas imagens da mesma cena tomadas por duas câmeras dispostas em locais distintos, mas próximos. A detecção do movimento, citada anteriormente, procura os mesmos elementos numa sequência de imagens tiradas ao longo do tempo, tentando estabelecer uma correlação entre seus posicionamentos consecutivos. Os desafios de utilizar a estereoscopia e a detecção do movimento são similares, pois ambos, essencialmente, tratam de identificar o mesmo objeto em duas ou mais imagens.

A implementação da estereoscopia computadorizada requer três condições. Em primeiro lugar, cada uma das câmeras estereo precisa ser equipada com um par de lentes delgadas. Em seguida vem que os eixos focais de ambas as lentes têm de ser paralelos entre si e perpendiculares ao segmento de base. E, finalmente, os planos focais de ambas as lentes são colineares, o que implica que ambos pertencem a um único plano. O cumprimento destes quesitos conduz à possibilidade do descarte do terceiro eixo, z , na análise da incoincidência. Girando adequadamente o sistema de coordenadas x - y - z , o pesquisador consegue visualizar a configuração 3-D numa representação 2-D.

Uma das principais vantagens inerente ao uso da estereovisão é a possibilidade de relacionar a incoincidência obtida a partir das duas imagens tomadas pelas lentes esquerda e direita com a distância entre o objeto e o segmento de base da câmera. Quando o objeto se encontra fora do plano focal (uma região retangular compreendida entre as duas lentes), a incoincidência é igual, essencialmente, à diferença entre a localização de um mesmo ponto da cena nas imagens das duas lentes. Muitos pesquisadores têm concluído que a distância é inversamente proporcional à incoincidência. Entretanto, esta afirmação é correta somente quando o ponto observado está localizado fora da região focal.

Quando o objeto está contido no plano focal, a distância e a incoincidência não são inversamente proporcionais entre si. A fim de minimizar o impacto desta discrepância,

o segmento de base da estereo câmera deve ser bastante reduzido. Uma linha de base pequena também facilita a estimativa dos valores da incoincidência porque aumenta a chance de que as duas lentes estejam registrando a mesma cena.



(Encima) A estereo câmera está fixada na lateral da coluna de sustentação do semáforo. Foto: Migma Systems, Inc.

(Embaixo) Os pesquisadores do FHWA montaram a estereo câmera sobre a coluna semafórica desta interseção em Norwood, MA. Foto: Migma Systems, Inc.

A questão que resta, agora, consiste em saber como calcular o valor da incoincidência. Em um esquema usual de estimação desses valores, um dos principais problemas é determinar os pontos conjugados em ambas as imagens, ou seja, descobrir os pontos das duas imagens que correspondem ao mesmo ponto real. A investigação toma por base um processo de comparação que avalia a similaridade entre pontos das duas imagens. Os pesquisadores, geralmente, utilizam uma das três seguintes abordagens para caracterizar a similaridade: a correlação, a intensidade e a fase.

A abordagem da correlação requer uma análise extremamente avançada da imagem a fim de conseguir identificar padrões equivalentes através do tratamento dos pixels. Pode ser mais confiável do que o método baseado na intensidade, especialmente para distâncias mais elevadas. Esta abordagem também sofre os problemas do alto volume de processamento e de problemas de classificação.

A abordagem da intensidade não requer a identificação das formas dos elementos, na medida em que todos os pixels são tratados da mesma forma. Sem contar a técnica da comparação entre blocos de elementos, os métodos fundamentados na análise da intensidade, através da utilização de métodos determinísticos ou estatísticos, têm sido predominantes no cálculo da incoincidência.

Ao longo da última década, muitos pesquisadores têm considerado a abordagem da fase como um dos mais eficientes métodos para estimar os valores da incoincidência. Este método se apóia no fato de que a informação da profundidade das imagens puras das lentes direita e esquerda pode ser relacionada à diferença da fase local entre ambas.

No sistema de estereovisão do FHWA, a incoincidência é estimada segundo a abordagem da correlação.



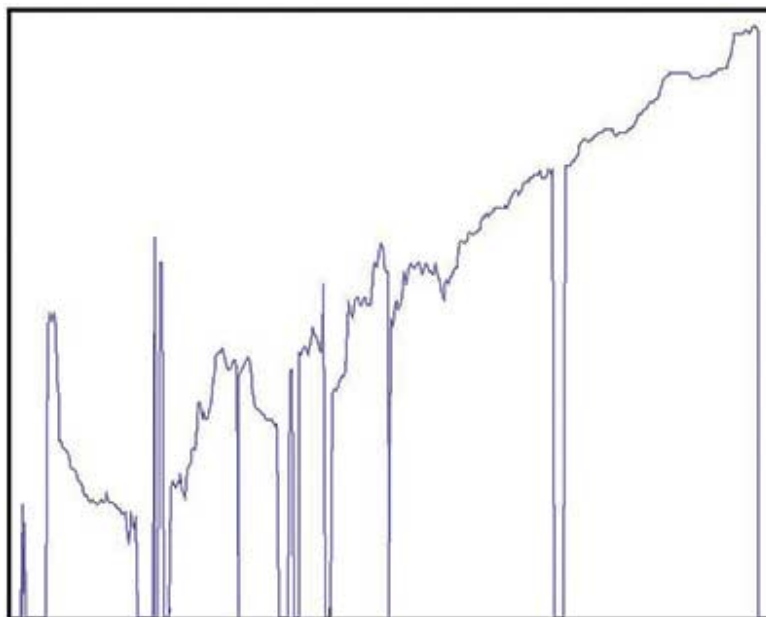
O mapa de incoincidência mostra um pedestre no centro da imagem. Como muitas vezes ocorre neste tipo de

representação, as pernas desaparecem ao se confundir com o solo. As lacunas (partes brancas) representam valores de incoincidência inválidos, o que significa que não foi possível obter nenhuma informação confiável a respeito da incoincidência nestas regiões.

O Sistema de Estereovisão do FHWA

O Sistema de Estereovisão do FHWA especifica uma câmera que possui duas lentes (esquerda e direita) com uma resolução de 1.024 por 768 pixels, um segmento de base de 12 centímetros e um ângulo horizontal de 70º do campo de visão. Como a câmera não possui alimentação elétrica própria, os pesquisadores a conectaram a um laptop, que lhe fornece a energia necessária. Foi aproveitado o software de interface da câmera, desenvolvido pelo seu próprio fabricante, para coletar os dados e estimar o mapa de incoincidência. Vários parâmetros da estratégia estereo, tais como resolução de imagem, tamanho da máscara, validação e distâncias focais são utilizados para calcular a incoincidência. Os pesquisadores utilizaram um programa especialmente desenvolvido a fim de ajustar os valores destes parâmetros às imagens tomadas. A câmera foi posicionada no alto de um poste em alturas que variaram entre 2,7 e 7,3 metros.

No sistema do FHWA, a frequência de coleta de dados variou entre 5 a 15 imagens (frames) por segundo. A frequência real, entretanto, depende da taxa de transmissão dos dados e das características da plataforma do computador. Os pesquisadores especificaram uma frequência real de 5 imagens por segundo (5 fps). Após analisar a velocidade média da travessia do pedestre, decidiram adotar o valor de 12 segundos para o período de amostragem, o que equivale a registrar 60 imagens para cada amostra. A seguir, utilizaram o sistema de estereovisão para registrar as travessias dos pedestres nos cruzamentos. As imagens coletadas foram classificadas em seis conjuntos diferentes. Foram consideradas três condições de iluminação ambiente: dia ensolarado, dia parcialmente nublado e dia nublado, com duas situações possíveis para cada um: travessia de um só pedestre ou de um grupo de pedestres. (Nota: os dados assim coletados podem ser requisitados no FHWA)



A figura mostra a variação da incoincidência ao longo de uma coluna vertical da imagem. O eixo y representa o valor da incoincidência enquanto que o eixo x mostra a correspondente coordenada vertical, assumindo-se o solo como origem. Os pesquisadores observaram que os valores da incoincidência crescem com a altura da coluna (embora não monotonicamente). Desta forma eles podem detectar os pedestres em layers específicos do diagrama de incoincidência. Fonte: Migma Systems, Inc.

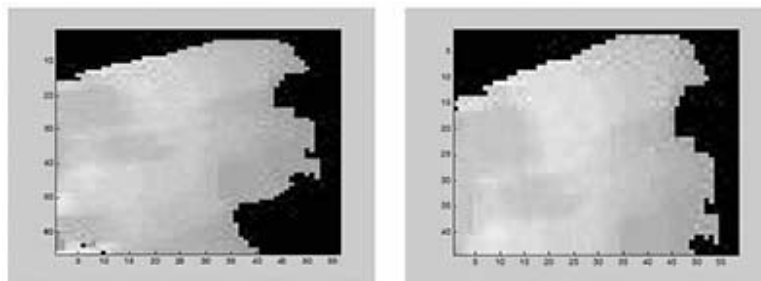
Valores limites do entorno

Um recurso gráfico denominado mapa de incoincidência é freqüentemente utilizado nos sistema de estereovisão. Geralmente, o tamanho deste mapa é igual ao tamanho tanto da imagem esquerda como direita. Cada pixel representa o valor da incoincidência no local correspondente e não o valor da intensidade da imagem naquele ponto. Como o valor da incoincidência está de alguma forma associado à distância entre o objeto e o segmento de base da câmera, o mapa de incoincidência pode ser visualizado como uma imagem 3-D. Lacunas ou regiões em branco neste mapa podem ser entendidas como zonas em que os valores de incoincidência encontrados foram inválidos, o que significa dizer que não existem tais valores para os correspondentes pixels.

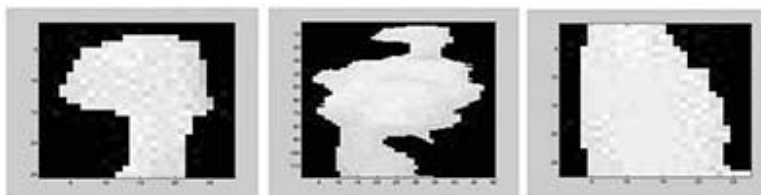
Num mapa de incoincidência, alguns elementos do entorno mais próximos da silhueta do pedestre, tais como o pavimento próximo aos seus pés, estão a aproximadamente a mesma distância da câmera. Isto acarreta que os valores de incoincidência serão similares e que é difícil separar os pixels associados com os pés do pedestre daqueles que correspondem com o solo próximo recorrendo tão somente aos valores de incoincidência. As formas associadas com os pedestres, no mapa, podem variar das originais encontradas nas imagens direita e esquerda. De fato, a forma do pedestre no

mapa depende fortemente da maneira como é calculado o fator incoincidência, o que implica, por exemplo, que a textura das roupas do transeunte pode vir a alterar a sua forma no mapa incoincidência. Por causa disso, a abordagem tradicional da detecção do pedestre através da comparação com padrões pré-determinados é menos confiável quando aplicada ao mapa de incoincidência.

Para superar este problema, os pesquisadores desenvolveram um método baseado na estimativa dos valores limites do entorno. Devido ao fato de que a os valores da incoincidência são normalmente (mas não sempre) inversamente proporcionais à distância, eles presumiram que tais valores deveriam aumentar no mapa ao longo das linhas verticais. Nesta previsão, o eixo y representa os valores da incoincidência e o eixo x representa as coordenadas horizontais do mapa. Os pesquisadores observaram que os valores cresciam (embora não monotonicamente devido ao "ruído" presente no mapa); desta maneira, eles conseguiram analisar o diagrama decompondo-o em layers (cada um deles decorrente de um diferente valor limite do entorno) e detectar os pedestres em cada um deles.



A figura mostra duas pequenas partes do entorno, extraídas do mapa de incoincidência. Fonte: Migma Systems, Inc.



A figura mostra três partes do corpo de um pedestre, extraídas do mapa de incoincidência. Fonte: Migma Systems, Inc.

A fim de distinguir o pedestre do resto da imagem, os pesquisadores desenvolveram um algoritmo para estimar os valores limites do entorno. Eles estimaram os valores destes limites para cada mapa de incoincidência construído. Para cada mapa, eles aplicaram o método do operador morfológico, utilizado na técnica de processamento de imagens e capaz de analisar objetos estruturados, encontrar pequenos vazios e converter o layer de incoincidência resultante numa imagem binária. Nesta imagem

binária, eles aplicaram um modelo de regressão matemática a fim de construir linhas limites do entorno que formassem uma moldura em torno de cada pedestre em potencial no mapa de incoincidência. Com esta abordagem, os potenciais pedestres são gradualmente extraídos da imagem, ao longo da análise dos layers. O número de layers depende do mapa original. Em outras palavras, os pesquisadores não fixaram previamente o número de layers a ser utilizado em cada caso.

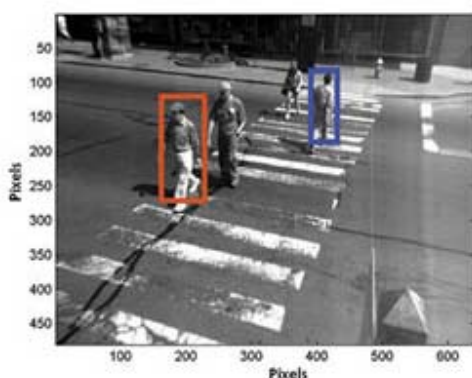


Imagem k1 com dois pedestres emoldurados

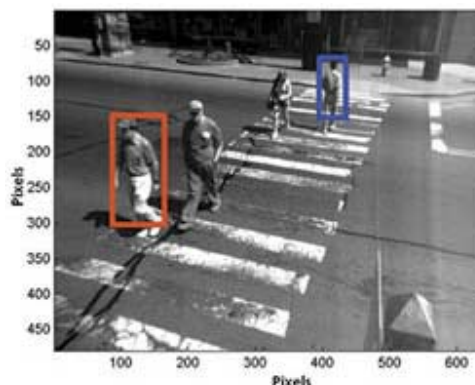
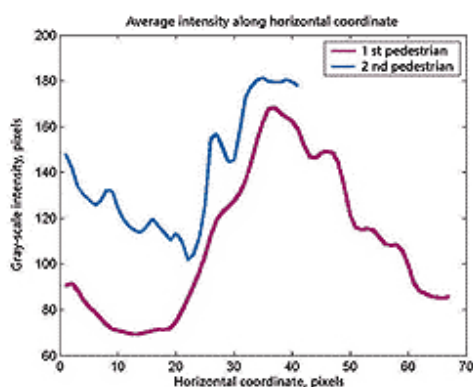
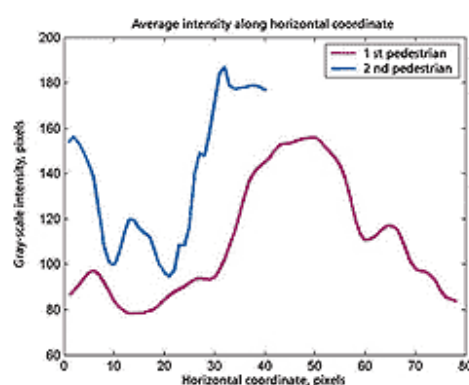


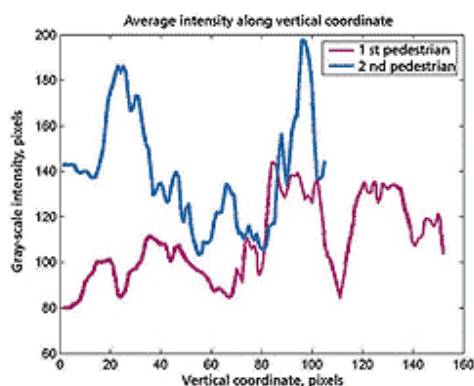
Imagem k2 com os mesmos dois pedestres emoldurados



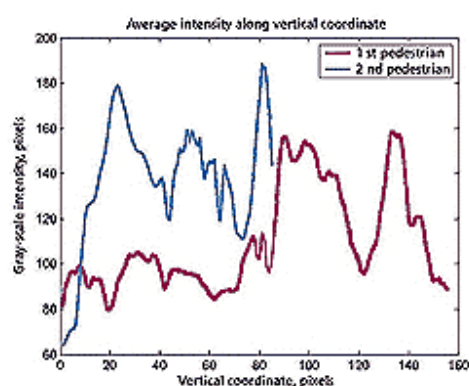
Valores das intensidades projetados no eixo X para os pedestres selecionados da imagem k1



Valores das intensidades projetados no eixo X para os pedestres selecionados da imagem k2



Valores das intensidades projetados no eixo Y para os pedestres



Valores das intensidades projetados no eixo Y para os pedestres

selecionados da imagem k1

selecionados da imagem k2

Estas figuras ilustram o princípio do rastreamento contínuo do pedestre. As duas imagens de vídeo (k1 e k2) foram tiradas com um intervalo de 1,2 segundos. As molduras azul e vermelha indicam os correspondentes pedestres, cuja forma aparece modificada de uma imagem para a outra. Entretanto, as curvas dos valores de intensidades projetados, mostradas nos campos abaixo, permanecem bastante parecidas, o que torna possível utilizá-las para determinar padrões característicos, ou seja, para reconhecer o mesmo pedestre em duas imagens. Além disso, a distância, em pixels, entre o mesmo pedestre em imagens seqüenciais é geralmente menor do que a distância entre diferentes pedestres, embora isso possa não ocorrer às vezes. O uso simultâneo destas duas propriedades permite tanto o rastreamento contínuo dos pedestres detectados como a eliminação dos objetos do entorno que poderiam vir a ser interpretados erroneamente como pedestres. Fonte: Migma Systems, Inc.

A detecção de pedestres utilizando a curvatura convexa em 3-D

Um problema comum associado com a abordagem de comparação com padrões pré-determinados reside na necessidade de construir uma biblioteca de padrões que possua um carácter suficientemente geral para poder lidar com as variações do mundo real, decorrentes de fenômenos de rotação, escala, brilho e ângulo de visão. A fim de superar esta dificuldade, os pesquisadores desenvolveram um novo método para detectar pedestres que toma por base o mesmo mapa de incoincidência. Ao invés de utilizar padrões pré-determinados, eles recorreram às características matemáticas relacionadas à curvatura convexa 3-D para detectar os pedestres em cada um dos layers do mapa.



As figuras à esquerda mostram duas imagens consecutivas

tiradas com um intervalo de 200 milissegundos, e, à direita, seus mapas de incoincidência correspondentes.

Neste método, o corpo humano não é tratado como um elemento único. Em vez disso, é dividido em partes características cujo reconhecimento é mais fácil. Mesmo assim, percebeu-se que não era também possível montar uma biblioteca com padrões pré-determinados que representassem todos os possíveis formatos das partes do corpo humano. A partir daí decidiram que uma abordagem distinta teria de ser adotada para conseguir distinguir estas partes no diagrama de incoincidência.

A nova abordagem dos pesquisadores do FHWA consistiu em trabalhar com características matemáticas da 3-D. Especificamente, partiram à procura de expressões matemáticas que traduzissem as curvaturas das partes do corpo humano desenhado no mapa de incoincidência. Diferentemente de outros objetos do ambiente viário, o corpo humano possui curvaturas em 3-D semelhantes a uma esfera (cabeça), a um plano (corpo) e a cilindros (pernas). Os pesquisadores utilizaram, então, as características de curvas 3-D para identificar o pedestre a partir de seus pedaços. Para cada coluna vertical, uma curva parabólica foi utilizada a fim de ajustar seus valores de incoincidência. Em seguida, os pesquisadores verificaram, para cada layer, se a curva ajustada era côncava ou convexa. Se uma parte do corpo do pedestre (como a cabeça ou a parte superior do corpo) era detectada, a mesma podia ser extraída, depois de enquadrá-la com uma moldura específica. Depois que toda a imagem tenha sido processada, todas as pequenas molduras foram combinadas para formar uma moldura maior que continha o pedestre inteiro.

Rastreamento contínuo do pedestre

O método de detecção descrito acima nem garante que os pedestres serão detectados na mesma ordem numa sequência de imagens de vídeo, nem garante que os resultados da análise estarão isentos de falsos alarmes (falsa detecção). O objetivo final do sistema é rastrear continuamente o pedestre e prever sua localização nos próximos instantes. Os pesquisadores, portanto, precisaram se assegurar que era possível conseguir uma detecção contínua e isenta de erros. Com o objetivo de identificar o pedestre e evitar os falsos alarmes, a equipe de pesquisadores se valeu de duas diferentes características: a configuração geométrica do local e a projeção média das intensidades das imagens ao longo dos dois eixos.

Os resultados práticos dos testes

Os pesquisadores testaram o método de detecção de pedestres com uma enorme quantidade de imagens tiradas com o sistema de estereovisão. As imagens foram capturadas em interseções reais durante as diferentes condições de dia ensolarado, parcialmente dublado e totalmente dublado. Foram registrados casos com pedestres isolados e com grupos de pedestres.

Os pesquisadores examinaram e analisaram as imagens e seus respectivos mapas de incoincidência. Por exemplo, eles olharam duas imagens tiradas em dias ensolarados, cada uma delas mostrando quatro pedestres atravessando na faixa. A partir do mapa de incoincidência, descobriram que era fácil identificar os dois primeiros pedestres que estavam mais próximos à câmera. Os outros dois pedestres, embora reconhecíveis no mapa de incoincidência, não tinham sua identificação tão nítida. Os pesquisadores julgaram, então, que o motivo estava na distância entre eles e a câmera, pois os últimos estavam a 6,4 metros enquanto que os primeiros estavam somente a 2,7 metros.

A fim de detectar todos os quatro pedestres, a equipe utilizou, inicialmente, a técnica dos valores limites do entorno a fim de obter um conjunto de layers de incoincidência. Detectaram, então, as silhuetas dos pedestres em potencial em cada layer através do método da análise das características da curvatura convexa em 3-D. Ao detectarem um pedestre num determinado layer, o mesmo era emoldurado, significando detecção. No final, todas as molduras foram combinadas para extrair o pedestre em inteiro. Os pesquisadores aplicaram, então, a classificação de Bayes, um método estatístico voltado para classificação de objetos, a fim de correlacionar os pedestres detectados nas imagens consecutivas. Tal abordagem conseguiu detectar com sucesso todos os quatro pedestres; os pesquisadores emolduraram cada um com uma diferente cor para facilitar o reconhecimento de cada um dos pedestres no conjunto de imagens consecutivas.



A abordagem conseguiu detectar corretamente todos os quatro pedestres. Cada um deles foi emoldurado com uma cor diferente, para facilitar seu reconhecimento nas imagens seqüenciais

Conclusão

Em resumo, o novo sistema de detecção de pedestres compreende quatro módulos básicos:

- (1) escolha de valores limite do entorno a partir dos quais é possível configurar layers de incoincidência; os pedestres em potencial serão extraídos layer por layer;
- (2) detecção de pedestres através das características matemáticas das curvaturas convexas das figuras em 3-D e conseqüente emolduramento de cada pedestre em cada layer;

(3) reconhecimento e acompanhamento do mesmo pedestre num conjunto de imagens seqüenciais;

(4) estimativa da velocidade do pedestre e sua localização.

A abrangência imaginada, no início dos trabalhos, no que se refere à coleta de dados, a fim de conseguir validar a abordagem da pesquisa, foi limitada somente às travessias de pedestres. Numa próxima etapa, prevê-se estender a coleta de dados para detectar e rastrear também os pedestres que se encontrarem próximos ao meio-fio, prestes a iniciar sua travessia.

Após testar o sistema utilizando as imagens reais tiradas em campo com o sistema de estereovisão, a equipe concluiu que os resultados comprovaram que o sistema tem potencial suficiente para ser utilizado nas aplicações do Intelligent Vehicle Initiative (IVI).

Eventualmente, os pesquisadores poderão vir a incorporar os softwares desenvolvidos em plataformas de hardware que poderão vir a ser instaladas nas interseções, utilizando técnicas de comunicação sem fio com os controladores de semáforos. Além disso, FHWA precisa aprimorar a tecnologia implementada a fim de melhorar a sua precisão e diminuir seus custos.

Once the detection system developed in this research is refined, State departments of transportation will be able to use it in their pedestrian safety programs. As South Carolina's Dodds says, "Systems that better enable pedestrians and motorists to share the pavement safely at the same time would be worthy of consideration."

Assim que o sistema de detecção desenvolvido nesta pesquisa for aperfeiçoado, os departamentos de trânsito poderão utilizá-lo em seus programas de segurança de pedestres. Como diz Dodds, da Carolina do Sul: "Sistemas que habilitam os pedestres e os motoristas a compartilhar o espaço viário de uma forma segura devem ser considerados muito atentamente".

* **David R. P. Gibson** integra a equipe de pesquisadores do Enabling Technologies Team do FHWA's Office of Operations Research and Development. É engenheiro civil formado no Virginia Polytechnic Institute and State University, onde fez mestrado em engenharia de trânsito. Suas áreas de atuação incluem tecnologia de sensoriamento, equipamentos de controle, modelagem e educação.

* **Bo Ling** recebeu seu título de mestrado em matemática aplicada em 1990 e de doutorado em engenharia elétrica em 1993 na Michigan State University. Ling tem trabalhado como pesquisador principal em numerosos projetos de pesquisa governamentais. Ele é co-fundador, presidente e gerente geral do Migma Systems, Inc. Tornou-se membro senior do Institute of Electrical and Electronics Engineers, Inc., em 1998 e leciona, em tempo parcial, no Department of Electrical and Computer Engineering at Northeastern University em Boston.

* **Michael Zeifman** formou-se em Ciência dos Materiais e Metalurgia no Leningrad Polytechnic Institute (atual St. Petersburg State Polytechnic University) na Rússia, em

1987. Recebeu o título de mestre em Garantia de Qualidade no Technion-Israel Institute of Technology em Israel em 1997 e o de doutor em Física da Confiabilidade, em 2000, também no mesmo instituto israelense. Após vários anos de pesquisas na Indústria e na Universidade, ele se ligou em 2005 à Migma Systems, Inc., como engenheiro pesquisador senior.

* **Shaoqiang Dong** Formou-se em 1994 e doutorou-se em 1997 na Xi'an Jiaotong University, China na área de Engenharia Mecânica. Ligou-se em 2004 à Migma Systems, Inc., como engenheiro de pesquisas.

* **Uma Venkataraman** formou-se em Ciência da Computação em 1996 na Madras University, India e passou vários anos trabalhando no desenvolvimento de programas de software, que é sua principal área de atuação. Entrou para a Migma Systems, Inc. como engenheira de software senior em 2003.