

Nova abordagem para reconhecimento biométrico baseado em características dinâmicas da íris humana

*Ronaldo Martins da Costa

ronaldomc12@gmail.com

*Adilson Gonzaga

*Escola de Engenharia de São Carlos

adilson@sel.eesc.usp.br

Abstract

Algoritmos que identificam pessoas através da biometria da Iris consideram as características de textura de uma imagem do olho humano. No entanto, todas as técnicas desenvolvidas até hoje, utilizam-se de características extraídas de imagens estáticas, ou seja, sem a comprovação se a mesma pertence a um ser vivo. Este trabalho propõe um método que além de avaliar as características de textura, avalie também características dinâmicas de contração e dilatação bem como o comportamento das características de textura em função dos reflexos naturais do olho. Com isso, é possível aumentar a segurança de dispositivos de reconhecimento através da íris, pois, somente íris vivas poderão ser utilizadas.

1. Introdução

Em 1936 um oftalmologista chamado Frank Burch sugeriu que as características da textura da íris humana poderiam ser utilizadas como método de identificação pessoal. Esta teoria foi documentada por James Daggarts em 1949. Em 1987 dois oftalmologistas, Aran Safir e Leonard Flom patentearam essa idéia e solicitaram a John Daugman, pesquisador do laboratório de computação da Universidade de Cambridge que criasse um algoritmo computacional capaz de realizar o reconhecimento. John Daugman criou e patenteou seu método em 1993 [1], [2]. Desde então a maior parte dos sistemas biométricos que trabalham com reconhecimento da íris baseiam-se no algoritmo criado por Daugman [3].

Atualmente as técnicas de identificação são extremamente precisas, realizando o processo completo em frações de segundo. Contudo, todos os métodos conhecidos na literatura são realizados sobre imagens estáticas. O reconhecimento é executado em apenas um frame ou uma imagem. Este processo é rápido e oferece a precisão necessária para identificar uma pessoa, mas é suscetível a fraudes.

Os métodos atuais não garantem que a pessoa esteja realmente presente ou se a imagem capturada é uma fotografia, prótese ou até mesmo um vídeo que de alguma forma possa ser acoplado ao dispositivo de captura tentando burlar o sistema.

2. Objetivos

O propósito deste trabalho é realizar a identificação pelas características dinâmicas do olho humano, analisando o comportamento das informações de textura e de outras características que se alteram em resposta a alteração da iluminação, levando em consideração padrões de variação ao longo do tempo.

Através de um protótipo especialmente construído para esta finalidade captura-se um vídeo com o registro das alterações da íris ao longo de um intervalo de tempo pré-definido.

O método proposto neste trabalho é capaz de detectar se a pessoa analisada está presente frente ao dispositivo de captura, além de poder também ser aplicado em oftalmologia para identificar anomalias nos reflexos, ou no próprio globo ocular/íris, muitas vezes causadas por doenças.

3. Metodologia

3.1 Captura das imagens da íris

A metodologia proposta neste trabalho é diferenciada desde a captura das imagens. Uma avaliação simplificada do funcionamento do sistema óptico humano fornece os subsídios necessários para tal.

O sistema óptico humano é composto basicamente pelo globo ocular e o nervo óptico [4]. A função do globo ocular é de capturar através dos cones e bastonetes os estímulos gerados pela luz. Estes estímulos são levados até o cérebro pelo nervo óptico que são então processados.

Na transmissão dos estímulos pelo nervo óptico, estes passam por uma região denominada quiasma

óptico conforme mostrado na figura 1. No quiasma óptico ocorre o cruzamento das fibras médias dos nervos ópticos, de tal forma que aproximadamente metade das fibras passa do nervo óptico direito para o nervo óptico esquerdo e vice-versa. Isso faz com que os olhos estejam “conectados”, e os reflexos a estímulos aplicados em um dos olhos sejam apresentados também no outro. Esta função fisiológica é denominada de “Reflexo Consensual”.

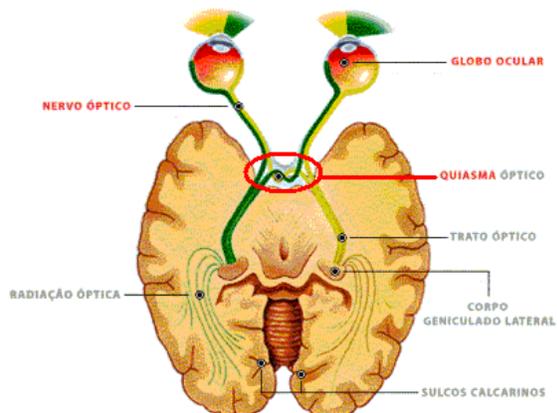


Figura 1: Sistema Óptico humano. Detalhe da região do quiasma óptico.

Baseado, então, no “Reflexo Consensual”, foi desenvolvido um equipamento de aquisição da imagem da íris apresentado na figura 2. O olho direito recebe estímulos de luz branca em tempos determinados controlados pelo software e detalhados na figura 3.

O olho esquerdo captura a imagem através de uma câmera e iluminação NIR (Near-Infra-Red). Como o olho não responde a esta banda de frequência, a abertura e fechamento da pupila do olho esquerdo são controlados apenas pela iluminação do olho direito através do “Reflexo Consensual”.



Figura 2: (a) Protótipo para captura dos vídeos
(b) Protótipo com a câmera “desmontada”

Cada vídeo é capturado com 1000 frames a uma taxa de 30 frames/segundo. No frame 210, um led com luz branca acende iluminando o olho direito, a luz permanece acesa até o frame 419. No frame 630 ela torna a acender e permanece acesa até o frame 839. Do frame 850 ao frame 1000, a luz é novamente apagada.

Esta mudança de iluminação faz com que a pupila contraia e dilate alterando tanto o seu diâmetro como a largura do anel formado pela íris.

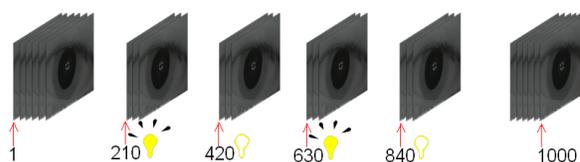


Figura 3: Estímulos de luz branca durante a aquisição das imagens iluminadas em infravermelho próximo.

Enquanto a iluminação é alterada no olho esquerdo, sincronamente a imagem de vídeo do olho direito é capturada.

No caso de uma implementação que vise sistemas de segurança, os intervalos (1 – 209; 210 – 419; 420 – 629; 630 – 839; 840 – 1000) podem ser alterados para tempos aleatórios, resultando em uma enorme dificuldade de fraude, pois em cada filmagem os tempos podem ser variáveis.

Para aplicação da abordagem proposta foram capturados vídeos de 66 pessoas, sendo 3 de cada uma, totalizando 198 vídeos.

3.2 Pré-processamento dos vídeos

Diferentemente de bases de imagens pré-preparadas e estáticas capturadas em ambientes controlados, a captura em tempo real, ou tempo de vídeo, apresenta maiores problemas. Os movimentos involuntários, ou não, são constantes, exigindo que a tarefa de pré-processamento seja bastante específica para descartar frames que possam comprometer o método e selecionar apenas frames que sejam adequados [5].

Em um mesmo vídeo podem existir frames ideais para o processamento, bem como frames desfocados, com problema de oclusão ou de angulação, conforme pode ser observado na figura 4.

Frames com problemas de oclusão ou desfocados devem ser desconsiderados na análise.

Quando o frame está em condição ideal para processamento, os movimentos da pupila oferecem detalhes para avaliação do comportamento do olho. A figura 5 apresenta um exemplo de frames ideais para o processamento.

No pré-processamento o algoritmo proposto equaliza cada um dos 1000 frames do vídeo.

Diversas sementes são “plantadas” em uma “janela” central do frame para realçar apenas a pupila e identificar seu centro. Calcula-se, então, o eixo maior, o eixo menor e a circularidade da pupila.

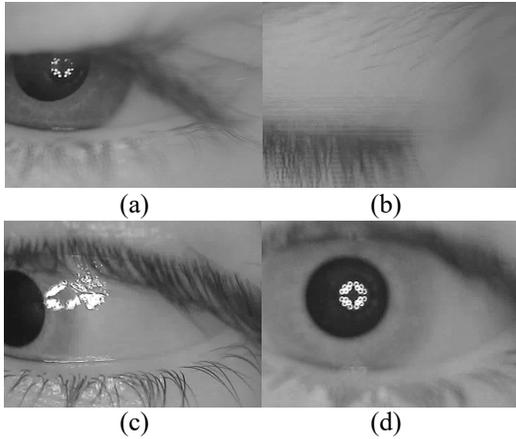


Figura 4: Exemplo de frames com problemas.

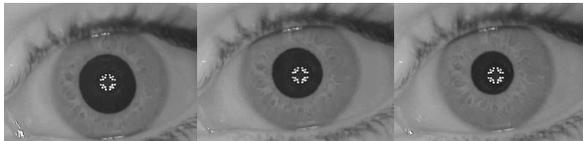


Figura 5: Exemplo de frames ideais para o processamento com movimentos de contração e dilatação da pupila.

O anel correspondente à íris é segmentado para análise das informações de textura. A parte superior e inferior do frame é descartada a partir do topo e do ponto inferior da pupila. A lateral esquerda e direita do frame é descartada através de um fator de multiplicação sobre o raio da pupila.

O resultado é uma imagem da íris segmentada conforme apresentada na figura 6.

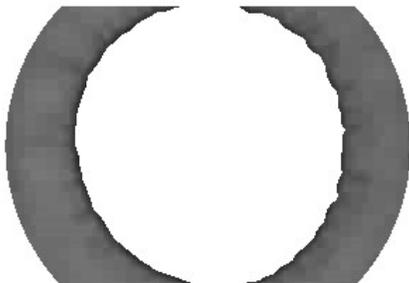


Figura 6: Exemplo de frame segmentado.

3.3 Extração das características

Para possibilitar a avaliação do comportamento de cada uma das características em função das alterações de iluminação, são estabelecidos cinco períodos de análise, sendo quatro períodos de transição, dentre os 1000 frames:

1º período – média geral de todos os 1000 frames;

2º período – média entre os frames 210 a 220;

3º período – média entre os frames 420 a 430;

4º período – média entre os frames 630 a 640;

5º período – média entre os frames 840 a 850;

Do 2º ao 5º períodos ocorrem alterações controladas de iluminação e a correspondente reação e adaptação do olho.

As características dinâmicas da pupila selecionadas para análise de comportamento padrão do olho são:

- + Circularidade;
- + Diâmetro;
- + Tempo para contração/dilatação;
- + Taxa de contração/dilatação;

São extraídas dez características de circularidade e diâmetro, duas de cada um dos cinco períodos.

O tempo e a taxa para contração/dilatação juntos correspondem a mais oito características do vetor.

Características estatísticas dos níveis de cinza da íris são também extraídas para compor o vetor. As características selecionadas são:

- + Média;
- + Desvio padrão;
- + Coeficiente de variação;

O semi-círculo formado pela íris segmentada é dividido em 2 partes (esquerda e direita) em função do corte superior e inferior realizado na segmentação. Cada uma dessas partes, denominadas aqui de Setor A e Setor B (setor A – lado esquerdo do semi-círculo e setor B – lado direito) fornecem mais trinta valores (3 características x 5 períodos x 2 setores) para composição do vetor de características.

A textura da íris tem sido utilizada tradicionalmente na identificação com imagens estáticas. Como a pupila se dilata/contrai nos períodos considerados, estas características apresentam comportamento padrão dinâmico, durante o movimento. São assim considerados os seguintes parâmetros de textura:

- + Correlação;
- + Segundo Momento Angular (SMA);
- + Entropia;
- + Contraste;
- + Momento da Diferença Inverso (MDI);

As características de textura [6] são obtidas sobre as matrizes de co-ocorrências geradas pela variação dos níveis de cinza nas direções de 0°, 45°, 90° e 135°.

As cinco características de textura geram mais 200 valores no vetor (5 características x 5 períodos x 2 setores x 4 ângulos).

3.3.1 Circularidade

Na figura 7 está destacado o contorno da pupila para o cálculo da circularidade.

A circularidade é uma característica dinâmica da pupila. Através do “Reflexo Consensual”, quando se aplica um pulso de luz no espectro visível (neste caso, luz branca) no olho direito, as duas pupilas contraem-se e quando a luz é retirada, as duas pupilas dilatam-se. A circularidade da pupila sofre alteração durante o tempo de contração e de dilatação que é medida frame a frame durante os períodos em questão (do 2º ao 5º). Além disso, as pessoas possuem circularidades diferentes o que representa uma característica individual.



Figura 7: Circularidade da pupila.

3.3.2 Diâmetro

O diâmetro da pupila é outra característica dinâmica que é utilizada neste trabalho. Como pode ser observado na figura 8, a pupila do exemplo contrai-se mais no primeiro pulso de luz, entre os frames 210 e 220 (2º período), do que entre os frames 630 e 640 (4º período). As médias individuais desta variação são utilizadas como característica de identificação.

3.3.3 Tempo de contração/dilatação

Após a alteração de iluminação o diâmetro da pupila chega a sua contração máxima ou mínima e depois se estabiliza. A região ressaltada no gráfico da figura 8 mostra a alteração do diâmetro na região de transição (pulso de luz aplicado no frame 210), o diâmetro mínimo atingido (contração máxima) e depois o retorno gradual ao valor médio estabilizado.

O tempo para contração/dilatação é calculado em número de frames, pois a taxa de aquisição da câmera de vídeo é constante (30 fps).

A quantidade de frames necessários para atingir o valor mínimo/máximo fornece o tempo para contração/dilatação do diâmetro da pupila.

Foi estabelecida a quantidade de 50 frames como limite máximo para avaliação, evitando-se assim erros gerados durante a fase de aquisição devido à reação da pessoa à luz branca aplicada.

3.3.4 Taxa de contração/dilatação

A taxa de contração e de dilatação da pupila é também discriminante para diferentes indivíduos.

A taxa de contração/dilatação é a derivada da variação do diâmetro dentro do período considerado, entre um valor mínimo e um valor máximo. Esta taxa é uma característica dinâmica que muda de indivíduo para indivíduo e é utilizada pela abordagem proposta, como característica de identificação.



Figura 8: Diâmetro médio da pupila.

3.3.5 Média aritmética

A média aritmética representa o valor médio dos níveis de cinza da íris nos setores e períodos selecionados. É importante salientar que os níveis de cinza não sofrem influência da iluminação, pois, a imagem é adquirida com luz infravermelha não causando reflexos e nem reações da pupila.

3.3.6 Desvio padrão

O desvio padrão é o desvio dos níveis de cinza da íris em torno da média, nos setores e períodos estabelecidos.

3.3.7 Coeficiente de variação

O coeficiente de variação é dado pela equação 1 e retrata o coeficiente de variação dos níveis de cinza da íris nos setores e períodos selecionados.

$$CV = \frac{S}{\bar{x}} \quad (1)$$

onde: S é o desvio padrão;
 \bar{x} é a média;

3.3.8 Correlação

A correlação é dada pela equação 2.

Ela mede a dependência linear entre os níveis de cinza de pares de pixels [6]. Valores próximos a um implicam numa forte relação entre os níveis de cinza dos pixels. A medida de correlação não é correlacionada com o SMA, isto é, altos valores de correlação podem ser encontrados em baixos ou altos valores de energia, para a mesma área de interesse.

$$cor = \frac{\sum_{i=1}^n \sum_{j=1}^n \frac{i \cdot j \cdot p(i, j, d, \theta)}{i \cdot j} - \mu_i \cdot \mu_j}{\sigma_i \cdot \sigma_j} \quad (2)$$

onde: i é o total de linhas da matriz.

j é o total de colunas da matriz.

$p(i, j, d, \theta)$ é o valor da célula de linha i , coluna j distância d e ângulo θ .

$$\mu_i = \sum_{i=1}^n \sum_{j=1}^n \frac{i \cdot p(i, j, d, \theta)}{i \cdot j}$$

$$\mu_j = \sum_{i=1}^n \sum_{j=1}^n \frac{j \cdot p(i, j, d, \theta)}{i \cdot j}$$

$$\sigma_i = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \frac{i^2 \cdot p(i, j, d, \theta)}{i \cdot j} - \mu_i^2}$$

$$\sigma_j = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \frac{j^2 \cdot p(i, j, d, \theta)}{i \cdot j} - \mu_j^2}$$

3.3.9 Segundo Momento Angular (SMA)

Esta medida avalia a uniformidade textural, que é a repetição de pares de níveis de cinza [6]. Quando a área de interesse apresenta textura uniforme (valores de níveis de cinza próximos) o valor de energia tende para 1. Caso a área não seja uniforme o valor da energia tende a 0 (zero). O SMA é dado pela equação 3.

$$SMA = \sum_{i=1}^n \sum_{j=1}^n \left(\frac{p(i, j, d, \theta)}{i \cdot j} \right)^2 \quad (3)$$

3.3.10 Entropia

A entropia mede a desordem em uma imagem [5]. Quando a imagem não apresenta textura uniforme, os valores da entropia tendem a ser muito baixos. A entropia alcança seu valor máximo quando os pixels na área de interesse apresentam níveis de cinza com valores aleatórios. A entropia apresenta uma correlação

linear negativa com o SMA e não é correlacionada com a medida de correlação. A entropia é dada pela equação 4.

$$ENT = \sum_{i=1}^n \sum_{j=1}^n \frac{p(i, j, d, \theta)}{i \cdot j} \cdot \log \left(\frac{i \cdot j}{p(i, j, d, \theta)} \right) \quad (4)$$

3.3.11 Contraste

O contraste mede a presença de transição abrupta de níveis de cinza, isto é, as bordas. Baixos valores indicam a ausência de bordas na área de interesse. Neste caso, a matriz de co-ocorrência de níveis de cinza apresenta valores concentrados em torno de sua diagonal principal. O contraste é calculado pela equação 5.

$$con = \sum_{i=1}^n \sum_{j=1}^n \frac{(i-j)^2 \cdot p(i, j, d, \theta)}{i \cdot j} \quad (5)$$

3.3.12 Momento da Diferença Inverso (MDI)

Quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima, o MDI atinge o valor máximo [7]. O MDI é calculado pela equação 6.

$$mdi = \sum_{i=1}^n \sum_{j=1}^n \frac{1}{1 + (i-j)^2} \cdot p(i, j, d, \theta) \quad (6)$$

4. Resultados

Para localizar um indivíduo foi gerado um vetor médio. O vetor médio é calculado pela média aritmética das características dos 3 vídeos reais.

O algoritmo gera um vetor temporário de 1000 características. Cada uma destas 1000 características é o resultado da média aritmética dos vídeos reais. Deste vetor temporário é extraído o vetor de características médio, conforme os períodos e setores citados no item 3.3 deste trabalho.

Para cada indivíduo foram adquiridas 3 seqüências de vídeos gerando-se 4 vetores de características, sendo três extraídos diretamente dos vídeos e um que é a média dos três vídeos.

Cada vetor de característica é formado por 12 valores: Circularidade, Diâmetro, Média, Tempo de Contração/Dilatação, Taxa de Contração/Dilatação, Desvio Padrão, Coeficiente de Variação, Correlação, SMA, Entropia, Contraste e MDI. Os dados extraídos

contêm as informações dos 1000 frames com seus respectivos setores (A, B) e graus (0°, 45°, 90° e 135°).

A figura 9 apresenta um gráfico de similaridades do vídeo 1 do indivíduo 1 com os vídeos da base.

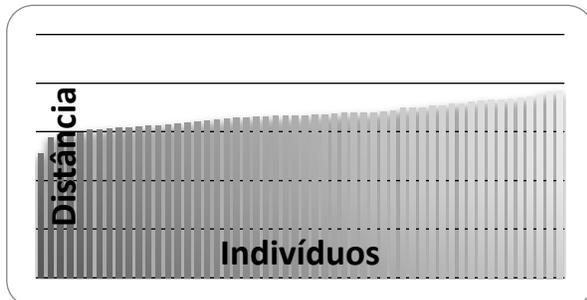


Figura 9: Gráfico de similaridade para o Vídeo 1 do Indivíduo 1

Neste exemplo, o vídeo 1 do indivíduo 1 foi comparado com todos os vídeos médios dos 55 indivíduos que compõem a base de vídeos, inclusive o do indivíduo 1.

É possível observar que a menor distância encontrada para o vídeo 1 do indivíduo 1 foi em relação ao seu próprio vetor médio.

O gráfico apresentado na figura 9 demonstra que é possível identificar um indivíduo através da avaliação do comportamento da íris em função de estímulos pré-determinados.

Ao todo foram feitos 161 vídeos. Destes, 47 foram descartados, por apresentarem problemas de iluminação ou foco. O algoritmo de extração de características foi projetado especialmente para eliminar estas medições. Vídeos com vetores de características com mais de 4 (quatro) características nulas eram automaticamente eliminados.

A figura 10 apresenta o gráfico que demonstra o percentual de acertos de identificação de 83,33% para os 114 vídeos analisados.

Os 16,67% de vídeos não localizados referem-se aos vídeos que não foram eliminados automaticamente pelo algoritmo, mas cuja qualidade não é ideal para análise.



Figura 10: Gráfico com o percentual de vídeos localizados e não localizados

5. Conclusões

É possível identificar biometricamente um indivíduo pelas características dinâmicas da íris. Através de um método de identificação mais robusto, utilizando o “Reflexo Consensual”, pode-se interagir o software de captura com o processo de identificação biométrico e assim verificar se a pessoa está realmente presente ou se é algum tipo de subterfúgio para burlar um sistema de segurança.

No entanto, como a proposta é orientada ao hardware, os resultados da identificação estão intimamente ligados a ele. Um hardware lento exige maior tempo de exposição e conseqüentemente maior desconforto ao usuário. As alterações de iluminação em geral fazem com que o indivíduo feche os olhos, prejudicando a aquisição dos vídeos. Um hardware mais aperfeiçoado pode capturar mais detalhes do movimento e também fornecer o resultado em um tempo menor.

Contudo, os resultados demonstram que a abordagem proposta, mesmo em condições não ideais de aquisição das imagens, é capaz de identificar um indivíduo com um alto índice de precisão, ou seja, é possível identificar um indivíduo pelas características dinâmicas de movimentos da íris em função de estímulos. As respostas aos estímulos seguem padrões individuais que devem ser mais investigados no sentido de se aumentar a precisão do sistema proposto.

Agradecimentos

Os autores agradecem à FAPESP – Fundação de Amparo à Pesquisa do estado de São Paulo, pelo suporte financeiro ao desenvolvimento deste trabalho de pesquisa.

Referências bibliográficas

- [1] Narote, S.P., Narote, A.S., Waghmare, L.M., Kokare, M.B., Gaikwad, A.N., An Iris Recognition Based on Dual Tree Complex Wavelet Transform, Pune Institute of Computer Technology, Pune, India, IEEE, 2007.
- [2] Daugman, J. G., *How iris recognition works*. Proceeding of 2002 International Conference on Image Processing, vol 1, 2002.
- [3] Daugman, J. G., *New Methods in Iris Recognition*. IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, Vol. 37, no. 5, 2007.

[4] Lana-Peixoto, M. A.. *Neuro-Oftalmologia. Sistema Sensorial – Parte II*. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492002000200020>. Acesso em 2 jun. 2008.

[5] Gonzaga, A., Moreno, R.P., Biometric Method for Personal Authentication Based on Iris Texture Features. In: The 4th Iasted International Conference on Visualization, Imaging, and Image Processing, v. 1. p. 1-5, 2004.

[6] Haralick, R.M.; Shanmugan, M.K., *Computer classification of reservoir sandstones*. IEEE Transactions on Geoscience Electronics, v.11, no. 4, p.171-177, 1973.

[7] Martins, S. P., *Classificação de imagens textural de imagens radarsat-1 para discriminação de alvos agrícolas*. São José dos Campos, INPE, Instituto Nacional de Pesquisas Espaciais, Dissertação de Mestrado, 2005.