MULTIBIOMETRIA: UMA VISÃO GERAL E APLICADA A FACE E VOZ

Multibiometrics: An Overview and its Application to Face and Speaker Recognition

VIOLATO, Ricardo Paranhos Velloso
Fundação CPqD

ANGELONI, Marcus de Assis
Fundação CPqD

SIMÕES, Flávio Olmos
Fundação CPqD

ULIANI NETO, Mário
Fundação CPqD; Faculdade de Jaguariúna; Faculdade Politécnica de Campinas

PEREIRA, Tiago de Freitas
Fundação CPqD

Resumo: Esse trabalho apresenta uma visão geral sobre o tema multibiometria, destacando alguns aspectos de interesse específicos das biometrias de face e de voz. São descritas as diversas fontes de informação biométrica e de que formas essas informações podem ser combinadas para produzir um resultado final. Esse texto não busca fornecer detalhes matemáticos das técnicas mostradas, e, portanto, não pode ser usado como referência para implementação, mas sim servir de guia inicial sobre o tema.

Palavras-chave: Biometria; multibiometria; fusão; verificação de locutor; verificação de face.

Abstract: This work presents an overview about multibiometrics, highlighting some specific aspects of interests of face and voice biometrics. The various sources of biometric information are described, as well as how these evidences can be combined in order to produce a final result. This text is not supposed to provide mathematical details about the techniques and therefore is not intended to be used as an implementation reference, but rather to serve as an initial guide on the subject.

Key-words: Biometrics; Multibiometrics; fusion; speaker authentication; face authentication.

Introdução

A tarefa fundamental de um sistema de gerenciamento de identidades é a determinação (ou verificação) da identidade de uma pessoa (ou a identidade alegada). Tal tarefa pode ser necessária por diversas razões, mas o objetivo primário é, na maioria do casos, prevenir que um impostor acesse algum tipo de recurso protegido (um computador, uma conta de banco, uma instalação de uma empresa, um país etc.).

Métodos tradicionais para estabelecer a identidade de um indivíduo são baseados em algo que a pessoa sabe (por exemplo, senhas) e/ou algo que a pessoa possui (por exemplo, seu documento de identidade ou o cartão de crédito). Entretanto essas representações da identidade podem ser facilmente perdidas, compartilhadas, manipuladas ou roubadas, comprometendo a segurança.

A biometria é uma forma de identificação baseada em características físicas (impressão digital, íris etc) e/ou comportamentais (assinatura, modo de andar etc) de uma pessoa. Nesse contexto, um sistema biométrico de gerenciamento de identidades apresenta uma opção natural aos métodos tradicionais, podendo substituí-los ou complementá-los, nesse último caso formando um sistema chamado de multifator.

Os sistemas biométricos podem usar uma gama de características físicas ou comportamentais de um indivíduo, incluindo impressão digital, face, geometria da mão, geometria dos dedos, íris, retina, assinatura, modo de digitar, modo de andar, impressão da palma da mão, voz, orelha, veias da palma da mão, odor ou DNA [12]. Na literatura da área, tais características são chamadas de traço, indicador, identificador, modalidade biométrica ou simplesmente de biometria, sendo esta última a nomenclatura usada preferencialmente ao longo deste texto.

Apesar dos sistemas biométricos apresentarem suas próprias limitações (ruído nos dados adquiridos, variações intra-classe, similaridade inter-classe, não-

universalidade e ataques de impostores) [11], um traço biométrico não pode ser facilmente perdido, roubado ou compartilhado. Além disso, também aumentam a comodidade do usuário, aliviando a necessidade de criar e lembrar senhas e/ou de carregar algo (token, cartão etc.).

Muitas das limitações dos sistemas unibiométricos podem ser contornadas ou atenuadas com o emprego de mais de uma biometria, formando um sistema biométrico multimodal. Mas o uso de mais do que uma biometria não é a única forma de multibiometria. Outras abordagens incluem o uso de mais do que uma amostra (por exemplo, uma imagem de face frontal, um perfil esquerdo e um perfil direito), mais do que um sensor, mais do que um algoritmo e/ou mais do que uma unidade (por exemplo, olho direito e olho esquerdo). Ademais, a fusão de todas essas informações pode se dar em diversos níveis (dados brutos, vetores de atributos, pontuações, rank, e decisão final).

Esse documento visa explorar o tema da multibiometria em suas diversas facetas, delineando as estratégias aplicadas em cada caso, mas detalhando os aspectos de maior interesse para as biometrias de face e voz destinadas à verificação.

O texto está organizado da seguinte forma: na Seção 1, são apresentados os problemas associados às tecnologias biométricas; na Seção 2 é introduzida a multibiometria e como ela pode evitar alguns dos problemas dos sistema unibiométricos; e na Seção 3, são apresentados os níveis de fusão de múltiplas fontes de informação biométrica.

1. Biometria

Cada biometria possui suas vantagens e desvantagens ou prós e contras e, por isso, a escolha de uma biometria para uma aplicação em particular não depende apenas do seu desempenho em termos de taxas de erros. Em [6], são identificados sete fatores que influenciam a adequação de certa característica física e/ou comportamental a ser usada em uma aplicação biométrica:

a) Universalidade: todos os indivíduos que irão utilizar o sistema devem possuir a biometria;

b) Unicidade: a biometria deve ser suficientemente diferente entre os indivíduos que irão utilizar o sistema;

c) Permanência: a biometria deve ser suficientemente invariante ao longo de um período de tempo;

d) Mensurabilidade: deve ser possível adquirir e digitalizar a biometria usando dispositivos que não causem incômodo aos usuários;

e) Desempenho: a precisão do reconhecimento e os recursos necessários para atingir tal nível de precisão devem ser adequados às restrições da aplicação;

f) Aceitabilidade: os indivíduos que irão utilizar o sistema devem estar dispostos a apresentar sua característica biométrica ao sistema;

g) "Fraudabilidade"[1]: deve-se considerar a facilidade com que determinada biometria pode ser falsificada (biometrias físicas) ou imitada (biometrias comportamentais), no cenário da aplicação.

Nenhuma biometria é capaz de atender integralmente a todos esses requisitos em todas as aplicações imagináveis. Alguns desses fatores são bastante subjetivos ou fortemente dependentes da aplicação desejada. Para as biometrias de face e voz, esse é o caso do desempenho, da aceitabilidade e das fraudes.

A aceitabilidade da biometria, em geral, está muito relacionada às questões de privacidade que vem sendo levantadas nos últimos tempos com a expansão dos sistemas biométricos. As pessoas querem saber se os dados biométricos poderão ser utilizados para rastrear ou seguir uma pessoa sem que ela saiba; se alguma doença pode ser detectada a partir dos dados biométricos; se os dados biométricos serão usados somente com o propósito definido ou se pode ser usado com outros objetivos, sem que isso estivesse previamente definido; se a condição social e financeira de uma pessoa pode ser deduzida a partir dos dados biométricos; e quais as consequências do comprometimento de uma base de dados biométricos.

Além disso, apesar das vantagens descritas na Introdução, o uso da biometria apresenta limitações que impõem desafios ao desenvolvimento de sistemas de gerenciamento de identidades, especialmente em aplicações de larga escala. As

[1] Suscetibilidade a fraudes

próximas subseções descrevem essas limitações, com ênfase nas biometrias de face e de voz. Além disso são apresentadas outras vulnerabilidades e dificuldades enfrentadas por sistemas biométricos em geral.

1.1. Ruído nos dados

Dados ruidosos podem ser gerados em decorrência de sensores defeituosos, de baixa qualidade ou com manutenção ruim, de condições ambientais desfavoráveis e até mesmo devido a certa condição de saúde do usuário.

Para a biometria de voz, o ruído ambiente, por exemplo decorrente de uma gravação feita em uma rua movimentada, a voz de uma pessoa resfriada ou rouca e a utilização de microfones de baixa qualidade ou até mesmo com tecnologias diferentes (electret ou carbon-button) [8] podem ser vistos como fatores que inserem ruído nos dados e, portanto, possivelmente degradam o desempenho do sistema.

No caso de face, a qualidade da câmera utilizada, em termos de resolução, frequência de amostragem, “auto-foco” etc., e as condições de iluminação, como um ambiente muito escuro ou muito claro, ou uma iluminação inadequada produzem ruído nos dados, também com influência no desempenho do sistema.

O uso de dispositivos de captura (microfones e câmeras) diferentes é, na realidade, um problema que pode ser resolvido simplesmente impondo-se como requisito da aplicação certo aparelho de certa marca e fabricante. No entanto, uma solução como essa ficaria atrelada ao fornecedor do dispositivo e, caso este venha a ser substituído (por questões de custo, qualidade ou até mesmo pelo avanço da tecnologia dos sensores), seria necessário recadastrar todos os usuários do sistema.

Por isso o desenvolvimento de técnicas de extração de atributos e algoritmos de comparação que funcionem com diferentes sensores é fundamental para a operação do sistema, com grande impacto na usabilidade.

1.2. Variações Intra-Classe

Variações intra-classe podem ser causadas pela própria interação com o sensor ou por variações naturais do traço biométrico com o tempo ou com o estado psicológico do indivíduo, principalmente para as características comportamentais.

Essas variações podem ser tratadas armazenando-se múltiplas referências biométricas de cada usuário e através de mecanismos de atualização das referências biométricas.

No caso da voz, na interação com o microfone, o usuário pode falar mais próximo ou mais distante do microfone, e mais alto ou mais baixo. Também sabe-se que a voz de uma pessoa modifica-se com o passar do tempo, sendo imperativo mecanismos de atualização das referências biométricas. Além disso, a voz de indivíduo é fortemente influenciada pelo seu estado psicológico (descansado, estressado, embriagado etc.).

Para a biometria de face, a interação com a câmera pode produzir imagens de diferentes ângulos do rosto. Para estes casos, pode-se criar referências biométricas de várias poses do usuário, ou, em soluções mais elaboradas, modelos tridimensionais da face. Assim como a voz, a face das pessoas muda com o tempo, mas também através de outros fatores, como cabelo, barba, maquiagem etc., sendo novamente necessário mecanismos de atualização das referências biométricas.

1.3. Similaridade Inter-Classes

Do ponto de vista dos algoritmos de reconhecimento de padrões usados em sistemas biométricos, a similaridade inter-classes pode ser vista como uma sobreposição do espaço dos vetores de atributos de diferentes classes, ou, nesse caso, de diferentes indivíduos. Especialmente em sistemas de identificação, quanto mais indivíduos cadastrados no sistema, essa sobreposição será mais frequente e maior será a taxa de falsas aceitações. Por isso, para um certo conjunto de atributos e algoritmo de reconhecimento, há um limite superior para o número de usuários que podem ser cadastrados, em função da taxa máxima de falsa aceitação requerida pela aplicação.

Tanto a face como a voz não são considerados traços biométricos únicos, sendo que o exemplo mais claro são os gêmeos idênticos (para os quais o reconhecimento de face tem direcionado sua atenção para marcas faciais e singularidades, que podem diferenciá-los).

1.4. Não-universalidade

Para uma certa biometria, o sistema pode não ser capaz de capturar dados ou dados adequados de todos os indivíduos. Há, portanto, uma taxa de falha de cadastro, ou FTE (do inglês *Failure to Enroll*) associada a cada biometria.

Para a biometria de face, salvo raríssimas exceções de pessoas com deformidades graves, todos os indivíduos possuem dados adequados da biometria.

Já para a voz, há o caso de indivíduos mudos, para os quais não é possível capturar dados. Além disso, outros problemas associados à fala (gagueira, língua presa, doenças que interferem na produção da fala etc.) podem tornar os dados capturados inadequados para o reconhecimento do locutor.

1.5. Vulnerabilidades

Além das fraudes associadas especificamente à biometria, um sistema de autenticação biométrica está sujeito a ataques em diversos outros pontos de sua implementação, como identificado em [9]. A Figura 1 resume os ataques possíveis.

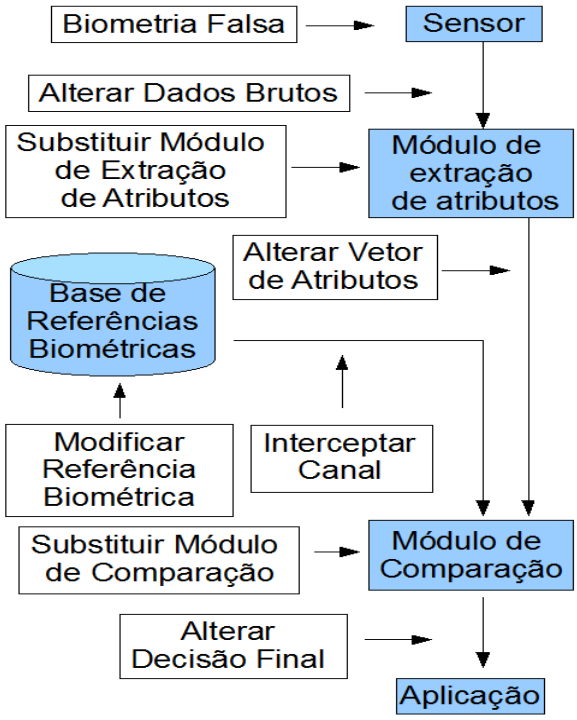

Figura 1: Pontos de Ataques em sistemas de autenticação biométrica (adaptada de [9]).

2. Multibiometria

Como descrito na Seção 1, o desenvolvimento e a aplicação de sistemas unibiométricos enfrentam vários empecilhos, muitos dos quais podem ser resolvidos ou amenizados com a multibiometria.

O termo multibiometria se refere a sistemas em que mais do que uma fonte de informação biométrica é utilizada na tarefa de reconhecimento. Essas múltiplas fontes de informação, ou de evidências, são então combinadas para produzir um resultado único. Há diversas formas de se combinar ou fundir as informações, como será visto na Seção 3.

A fusão de dados ou de informações é empregada em diversas áreas de pesquisa (mineração de dados, otimização, aprendizagem de máquina, robótica etc.), com diferentes terminologias, como, por exemplo, *ensembles*, *pool*, consenso etc. De forma geral, os sistemas desenvolvidos por essas áreas, inclusive a de biometria, podem ser chamados de sistemas de múltiplos classificadores.

Sistemas de múltiplos classificadores, ou MCS (do inglês *Multiple Classifier System*), buscam explorar as diferentes virtudes e escapar das diferentes fraquezas de diferentes classificadores, com objetivo de melhorar o desempenho global da aplicação. No caso da biometria, a palavra desempenho não está restrita ao sentido de precisão ou de taxas de erro do sistema, mas inclui fatores que serão listados a seguir.

Sistemas multibiométricos apresentam as seguintes vantagens em relação aos sistemas unibiométricos [11]:

- Sistemas multibiométricos podem ter desempenho significativamente superior, melhorando simultaneamente as taxas de falsa aceitação (FAR) e falsa rejeição (FRR);
- Em algumas situações, a disponibilidade de múltiplas fontes de informações aumenta o espaço dos atributos, aumentando assim também o número de indivíduos que podem ser cadastrados;
- O uso de múltiplas biometrias (um dos casos de sistema multibiométricos) aumenta a cobertura da população hábil a ser cadastrada;
- As possibilidades de fraudes biométricas são reduzidas. Além disso, permite maior liberdade na criação de mecanismos de desafio-resposta;

- Sistemas multibiométricos podem lidar melhor com o problema de ruído nos dados;
- Em situações de monitoramento ou rastreamento, sistemas multibiométricos permitem que o reconhecimento seja contínuo ou dure mais tempo;
- Um sistema multibiométrico pode ser visto como um sistema tolerante a falhas, uma vez que se certa fonte de informação biométrica falhar (um sensor, um software etc.) outras serão usadas.

As próximas subseções apresentam as muitas fontes de informação biométricas possíveis, de acordo com [11]. Além delas, há os sistemas híbridos, que integram mais do que um dos cenários que serão vistos.

Também não podemos nos esquecer de citar os sistemas multi-fator, em que a biometria é combinada com as formas tradicionais de autenticação, como senhas e tokens, aumentando a segurança.

2.1. Sistemas Multi-Amostras

Um único sensor pode ser usado para capturar múltiplas amostras da mesma biometria. Com isso, consegue-se descrever variações naturais do traço biométrico ou obter uma representação mais completa da biometria. Um aspecto importante neste caso é determinar quantas amostras devem ser capturadas, para que nelas esteja presente a variabilidade e a tipicidade da biometria naquele indivíduo.

Para a biometria de voz, pode-se dizer que os sistemas são inerentemente multi-amostras, já que o sinal de voz é sempre analisado por partes (dividido em quadros e, para cada quadro, é calculado um vetor de atributos). Quanto tempo um usuário deve falar e o que ele deve falar, para que suas amostras representem adequadamente a variação natural da sua voz, são questões que devem ser consideradas em um sistema de reconhecimento de locutor.

No caso da face, um sistema multi-amostras poderia, por exemplo, usar uma imagem frontal da face e imagens do perfil lateral de um indivíduo. Também poderia ser usado um vídeo, ou seja, diversas amostras da face, para a tarefa de reconhecimento.

2.2. Sistemas Multi-Sensor

Uma mesma biometria pode ser capturada por diferentes sensores, cada um com características específicas e produzindo informações diferentes.

No caso da biometria facial, há sensores para obter imagens tridimensionais, câmeras tradicionais e sensores para gerar imagens térmicas, sendo que a combinação desses dois últimos já se mostrou vantajosa [2].

Quanto à voz, sabe-se que diferentes tipos de microfone introduzem diferentes comportamentos no sinal capturado [8]. No entanto, geralmente essas variações são tratadas com técnicas de processamento de sinal ou de parâmetros.

2.3. Sistemas Multi-Algoritmo

Esses sistemas podem usar múltiplos algoritmos para extração de atributos ou para comparação, processando o mesmo dado biométrico. Contudo a adição desses módulos pode aumentar a complexidade computacional do sistema, exigindo mais recursos. Os fatores que podem influenciar um sistema multi-algoritmo incluem a correlação entre esses algoritmos, a disparidade de precisão das comparações e a metodologia de fusão adotada.

No campo do reconhecimento de face, há uma grande variedade de abordagens para a extração de atributos. Já no processamento de fala, a grande maioria dos trabalhos na área utiliza como atributos os coeficientes mel-cepstrais.

Quanto aos algoritmos de comparação, novamente a área de reconhecimento de locutor está limitada a apenas duas abordagens, GMM (do inglês *Gaussian Mixture Model*) e SVM (do inglês *Support Vectors Machines*). Enquanto isso, para a biometria de face há diversas estratégias, incluindo, além do GMM e do SVM, o OPF (do inglês *Optimum-Path Forest*), Máxima Verossimilhança (baseada na Regra de Decisão de Bayes), K-Vizinhos Mais Próximos, ou KNN (do inglês *K-Nearest Neighbors*), medidas de dissimilaridade mais simples (tais como Qui-quadrado, Intersecção de Histograma, Distância Euclidiana, Mahalanobis), entre outras.

 ISSN 1679-8902

2.4. Sistemas Multi-Unidade

São sistemas que usam mais do que uma instância do mesmo traço biométrico. Exemplos dessa abordagem são sistemas que utilizam mais de um dedo para impressão digital ou a íris dos dois olhos de um indivíduo. Esse caso não se aplica para face nem para voz.

Nesses sistemas não é necessário existir mais do que um sensor (apesar de poder ajudar na usabilidade) e também não são necessários novos algoritmos de extração de atributos ou de comparação. Além disso eles aumentam a cobertura da população, a capacidade de um sistema de identificação (número de usuários cadastrados) e, naturalmente, a precisão do sistema.

2.5. Sistemas Multimodais

Sistemas multimodais não se referem a uma biometria individualmente, mas a combinação de múltiplas biometrias. Um sistema biométrico multimodal poderia, por exemplo, requerer dados da voz e da face de um usuário.

Esses sistemas envolvem maior custo de desenvolvimento e da aplicação final, já que necessitam de diferentes sensores, diferentes algoritmos de extração de atributos e provavelmente diferentes algoritmos de comparação, além de poder ter um impacto negativo em termos de usabilidade.

No entanto, eles contemplam as vantagens das outras formas de multibiometria, isto é, têm um grande potencial de ganho de desempenho, eles aumentam a cobertura da população, dificultam fraudes biométricas, são mais robustos a ruídos, apresentam menor variação intra-classe e maior separação inter-classes e, especialmente para sistemas de identificação, podem armazenar mais usuários.

3. Níveis de Fusão

As diversas fontes de informação biométrica descritas na Seção 2 podem ser combinadas de diferentes formas, ou em diferentes estágios de um sistema multibiométrico, desde dados brutos até o resultado final.

Nesta seção, será descrita a fusão nos níveis de dados brutos, de atributos, de pontuação, de rank (para sistemas de identificação) e de decisão final (geralmente para sistemas de verificação). Nessa ordem, a quantidade de informação disponível diminui a cada etapa do processo de reconhecimento, partindo dos dados brutos, em que teoricamente está disponível toda a informação, até a decisão final, que pode ser representada por apenas 1 bit.

Deve-se atentar para o fato de que nem sempre a disponibilidade de mais do que uma fonte de informação biométrica implica que a fusão deve ser feita. Primeiramente, é necessário analisar a correlação entre essas informações. Combinar informações correlacionadas pode não trazer ganho algum, enquanto espera-se que a fusão de informações com correlação negativa ou descorrelacionadas produza maiores ganhos.

3.1. Fusão no Nível do Sensor

A fusão no nível do sensor consiste em combinar dados brutos de diferentes fontes de informação biométrica. Geralmente é aplicada em uma mesma biometria, obtida ou com sensores diferentes (multi-sensor) ou com várias amostras capturadas com o mesmo sensor (multi-amostras).

Uma técnica conhecida como combinação em mosaico (do inglês *mosaicing*) é aplicada para impressões digitais e faces, e consiste em uma combinação de múltiplas amostras. Esse é o caso dos sistemas de impressão digital que utilizam um sensor, no qual se rola o dedo, comum em algumas marcas de notebook. Para face, conforme mencionado na Seção 2.1, imagens de poses diferentes podem ser combinas em uma única imagem.

Além do caso de *mosaicing*, existem exemplos de sistemas multi-sensores com fusão no nível do sensor para biometria de face, os quais empregam uma combinação de imagens bi- e tridimensionais, citado também na Seção 2.2.

3.2 Fusão no Nível dos Atributos

Fusão no nível dos atributos significa combinar múltiplos vetores de atributos obtidos do mesmo indivíduo. Algumas dificuldades surgem ao se tentar implementar a fusão neste nível:

- A relação entre os espaços dos atributos pode ser desconhecida;
- Os múltiplos vetores de atributos podem ser incompatíveis. Por exemplo, cada elemento de vetores de minúcias de impressões digitais parametriza a posição relativa das minúcias e o vetor tem tamanho variável, de acordo com a quantidade de minúcias detectadas. Esses dados não poderiam ser combinados com um vetor de tamanho fixo, constituído de escalares, como é o caso dos auto-coeficientes de uma face;
- Ainda que os vetores sejam de mesmo tamanho, concatená-los pode aumentar muito a dimensão final do vetor de atributos, propiciando o surgimento do mal da dimensionalidade;
- Poucos sistemas biométricos comerciais fornecem acesso aos atributos usados, dificultando o progresso das pesquisas na área.

Duas possibilidades devem, então, ser consideradas.

Uma delas ocorre se os vetores de atributos foram obtidos pelo mesmo algoritmo de extração, e portanto advindos da mesma biometria, e a referência biométrica é o próprio vetor de atributos, como na biometria de impressão digital ou de geometria da mão, por exemplo. A outra situação ocorre quando os vetores de atributos foram gerados por algoritmos de extração diferentes, geralmente associados a biometrias diferentes.

No primeiro caso, a fusão no nível dos atributos pode ser usada para o aperfeiçoamento da referência biométrica (adição de características não observadas anteriormente e remoção de dados espúrios) ou para sua atualização (para acomodar variações temporais do traço biométrico).

Na segunda situação, os atributos devem ser primeiramente normalizados para que assumam valores em um mesmo intervalo válido e para que suas distribuições sejam também próximas. Em seguida, recomenda-se o uso de alguma técnica de redução de dimensionalidade, como a seleção de atributos ou a transformação de atributos.

Para a biometria de face, há exemplos de sistemas multi-algoritmo e multi-sensor que realizam a fusão no nível dos atributos [10]. Este aplica uma análise de discriminantes lineares, ou LDA (do inglês *Linear Discriminant Analysis*), independentemente em três canais de cores diferentes e combina os resultados. Naquele são combinados atributos de um algoritmo baseado em análise de componentes principais, ou PCA (do inglês *Principal Component Analysis*), com outro baseado em LDA.

Para a biometria de locutor, já foi proposto combinar os atributos da voz com o formato dos lábios [3].

3.3 Fusão no Nível do *Rank*

A fusão no nível do *rank* só se aplica a sistemas de identificação. A saída desses sistemas é uma ordenação, ou *ranking*, dos usuários cadastrados, em que a identidade que está sendo avaliada é associada mais provavelmente ao primeiro indivíduo da lista (*rank* 1) e assim por diante.

Há três métodos de combinar os *rankings* gerados por diferentes sistemas [4]:

Método do maior *rank*: como o próprio nome diz, cada usuário cadastrado recebe o maior *rank* fornecido pelos sistemas disponíveis. Nos casos em que ocorre empate (sistema A indica *rank* 1 para o usuário x e o sistema B indica *rank* 1 para o usuário y), o *rank* final é escolhido aleatoriamente (escolhe-se aleatoriamente se o *rank* 1 será o usuário x ou y).

Método da soma (*Borda Count Method*): nesse método, o *rank* de cada usuário é dado pela soma dos *ranks* obtidos por cada sistema.

Método da regressão logística: é uma generalização do método da soma, em que o *rank* gerado por cada sistema recebe um peso diferente, de acordo com a precisão de cada um deles. O peso é determinado por uma regressão logística e esse método depende de uma etapa de treinamento para ajustar os pesos.

 ISSN 1679-8902

3.4. Fusão no Nível da Decisão

Em sistemas de verificação, pode-se usar apenas o resultado final, ou a decisão, de diversos algoritmos de comparação, ou classificadores, e combiná-los para melhorar o desempenho do sistema como um todo. Além disso, a fusão no nível da decisão pode ser utilizada mesmo em um sistema de identificação, pois muitas vezes só é disponibilizada a classe vencedora e não o ranking completo.

Exemplos de técnicas simples empregadas para a fusão no nível da decisão são:

- Regra do “E” ou Regra do “OU”: esses são os métodos mais simples de fusão no nível da decisão. No caso da regra do “E”, a resposta do sistema confirmará a autenticidade do indivíduo somente se todos os algoritmos de comparação utilizados concordarem. No caso da regra do “OU” basta que apenas 1 dos classificadores confirme a autenticidade, mesmo que os outros neguem;
- Voto majoritário: este é provavelmente o método mais comum de fusão no nível da decisão e funciona da seguinte forma: a resposta do sistema é dada pelo resultado da maioria dos algoritmos de comparação;
- Voto majoritário ponderado: similar ao voto majoritário, mas sendo que o resultado de cada classificador recebe um peso, conforme sua precisão. A decisão final é dada em função da soma ponderada dos resultados de cada algoritmo.

Além dessas abordagens mais imediatas, algumas técnicas empregadas para a fusão no nível da decisão envolvem uma etapa de treinamento, na qual o sistema utiliza uma base de dados para “aprender” como fundir os resultados dos classificadores. Exemplos são a teoria de evidências de Dempster-Shafer [13], o espaço de conhecimento do comportamento, ou BKS (do inglês *Behavior Knowledge Space*) [5], e o método de decisão Bayesiano [13].

3.5. Fusão no Nível de Pontuação

A fusão no nível da pontuação ocorre quando as pontuações geradas por diferentes algoritmos de comparação são consolidadas para se produzir a decisão

final. A pontuação é uma medida de similaridade entre o vetor de atributos de entrada e uma referência biométrica.

Três fatores fazem com que a fusão no nível da pontuação seja uma tarefa desafiadora. Primeiro, alguns algoritmos de comparação podem produzir uma medida de distância ou de dissimilaridade, enquanto outros podem retornar uma medida de similaridade. Segundo, as pontuações podem estar em escalas numéricas e intervalos válidos diferentes. Por fim, as pontuações podem ter distribuições de probabilidade diferentes.

Do ponto de vista do reconhecimento de padrões, a tarefa de um sistema biométrico é designar um dado padrão de entrada a uma das possíveis classes disponíveis. Dado que, em um sistema multibiométrico com diferentes classificadores, são gerados diferentes vetores de atributos para cada classificador, deve-se atribuir o padrão de entrada àquela classe que maximiza a probabilidade conjunta dos diversos vetores de atributos.

Considerando que os vetores de atributos dos diferentes classificadores são estatisticamente independentes, em [7] foram propostas cinco estratégias, ou regras, de combinação de classificadores: regra do produto, regra da soma ou da média, regra do máximo, regra do mínimo e regra da mediana.

A hipótese de independência estatística é naturalmente válida em um sistema multimodal, mas, por exemplo, em um sistema multi-amostras, as diferentes amostras de uma mesma biometria tendem a ser muito correlacionadas.

Essas regras são aplicáveis somente se a saída dos algoritmos de comparação forem a probabilidade a posteriori de cada classe, dado o padrão de entrada. Geralmente, os classificadores retornam uma pontuação, que, de forma geral, podem ser consideradas como uma função da probabilidade adicionada de um erro. Foram então propostas técnicas para se estimar as probabilidades a posteriori, a partir das pontuações produzidas pelos algoritmos de comparação. Essas técnicas podem ser divididas em três grandes categorias [11]:

- Fusão de pontuações baseada em densidade: nessa abordagem, considera-se que as probabilidades a posteriori das classes, dados os vetores de atributo, podem ser diretamente aproximadas pelas probabilidades, dadas as pontuações. A conversão de pontuações em probabilidades requer

implicitamente a estimação das densidades condicionais da pontuação dada a classe, de onde vem o nome dessas abordagens;

- Fusão de pontuação baseada em transformação: estimar as densidades condicionais das pontuações, dada a classe, só é possível quando há uma quantidade grande de dados de treinamento. Além disso, as probabilidades a posteriori das classes, dados os vetores de atributo, normalmente não podem ser diretamente aproximadas pelas probabilidades, dadas as pontuações, pois deve-se somar o termo de erro de aproximação diferente de zero. Nessa abordagem, as pontuações geradas pelos diferentes algoritmos de comparação são transformadas para um espaço comum, tornando-as comparáveis. Essa abordagem é conhecida como normalização de pontuação e, geralmente, as pontuações normalizadas não tem interpretação probabilística;
- Fusão de pontuação baseada em classificador: nessa abordagem, a relação entre as pontuações geradas e as probabilidades a posteriori das classes, dadas as pontuações, são aprendidas por um algoritmo de reconhecimento de padrões.

Técnicas de normalização de pontuação são usadas em sistemas reconhecimento de locutor, mesmo sem a finalidade de fusão, pois muitas vezes facilitam a análise dos resultados e a escolha de limiares. As técnicas mais comuns são Z-norm e T-norm [1].

Além disso, a fusão no nível da pontuação é mais comum em sistemas multimodais. Por exemplo, o projeto MOBIO [2] tinha o objetivo de justamente desenvolver sistemas de reconhecimento de locutor e de face e, por conseguinte, de combiná-los.

De forma geral, todas as técnicas citadas podem ser aplicadas para ambas as biometrias, face e voz.

Um ponto a ser destacado é que, no caso dos sistemas de verificação, as técnicas de fusão no nível da pontuação são aplicados da mesma forma, sendo que existem apenas duas classes, “genuíno” ou “impostor”, ao contrário de sistemas de

[2] http://www.mobioproject.org/

identificação em que as classes são tantas quanto o número de usuários cadastrados.

Em um sistema multibiométrico, a pontuação fornece o melhor compromisso entre a quantidade de informação disponível e a facilidade com que é possível fazer a fusão. Por isso, a fusão no nível da pontuação é a técnica dominante nos sistemas multibiométricos, sendo a mais estudada e, portanto, para a qual há mais material disponível.

Considerações Finais

Nesse documento foi apresentada uma visão geral dos sistemas multibiométricos, ou seja, sistemas em que mais do uma fonte de informação biométrica é combinada para tarefa de reconhecimento.

Foi fornecida uma contextualização dos fatores que devem ser levados em conta na implantação de um sistema biométrico, mostrando os problemas e desvantagens relacionados a sistemas unibiométricos e como eles podem ser contornados, ou pelo menos atenuados, com sistemas multibiométricos.

Foram descritas essas fontes de evidências, multi-amostras, multi-sensor, multi-algoritmo, multi-unidade e multimodal, destacando-se fatores que se aplicam nas biometrias de voz e de face.

Também foram analisados de que modos as múltiplas fontes de informação biométrica podem ser combinadas, ou seja, os níveis de fusão: fusão no nível sensor, ou dos dados brutos, no nível dos atributos, no nível da pontuação, no nível do rank e no nível da decisão final. Novamente, foram salientados, quando necessário, os aspectos relativos as biometrias de face e de voz.

Esse documento não se aprofundou em detalhes matemáticos das técnicas apresentadas, haja vista que, como o título sugere, seu objetivo é fornecer uma visão geral sobre o tema multibiometria, levantando alguns pontos de interesse específicos para as biometrias de face e de voz.

Portanto, esse texto deve ser usado como material de consulta inicial e não como um guia para implementação das técnicas aqui citadas.

Bibliografia

[1] R. Auckenthaler, M. Carey e H. Lloyd-Thomas. **Score Normalization for Text-Independent Speaker Verification Systems**. Digital Signal Processing, *v*ol. 10, nos. 1–3, 2000.

[2] X. Chen, P. J. Flynn e K. W. Bowyer. **IR and Visible Light Face Recognition**. Computer Vision and Image Understanding, 99(3), pp. 332-358, 2005.

[3] C. C. Chibelushi, J. S. D. Mason e F. Deravi. **Feature-level Data Fusion for Bi-modal Person Recognition**. In *Proceedings of the Sixth International Conference on Image Processing and Its Applications*, vol. 1, pp. 399-403, Dublin, Ireland, 1997.

[4] T. K. Ho, J. J. Hull e S. N. Srihari. **Decision Combination in Multiple Classifier Systems**. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(1), pp. 66-75, 1994.*

[5] Y. S. Huang e C. Y. Suen. **Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals***. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(1), pp. 90-94, 1995.*

[6] A. K. Jain, R. Bolle e S. Pankanti (editors). **Biometrics: Personal Identification in Networked Society***. Kluwer Academic Publishers, 1999.*

[7] J. Kittler, M. Hatef, R. P. Duin e J. G. Matas. **On Combining Classifiers***. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3), pp. 226-239, 1998.*

[8] N. Malayath, H. Hermansky, S. Kajarekar e B. Yegnanarayana. **Data-Driven Temporal Filters and Alternatives to GMM in Speaker Verification**. In *Digital Signal Processing*, vol. 10 (1-3), pp. 55-74, 2000.

[9] N. K. Ratha, J. H. Connel e R. M. Bolle. **An Analysis of Minutiae Matching Strength.** In *Proceedings of Third International Conference on Audio- and Video-Based Biometric Person Authentication* (AVBPA), pp. 223-228, Halmstad, Suécia, 2001.

[10] A. Ross e R. Govindarajan. **Feature Level Fusion Using Hand and Face Biometrics**. In *Proceedings of SPIE Conference on Biometric Technology for Human Identification II*, vol. 5779, pp. 196-204, Orlando, USA, 2005.

[11] A. Ross, K. Nandakumar e A. K. Jain. **Handbook of Multibiometrics**. Springer, 2006.

[12] J. L. Wayman, A. K. Jain, D. Maltoni e D. Maio (editors). **Biometric Systems: Technology, Design and Performance Evaluation**. Springer, 2005.

[13] L. Xu, A. Krzyzak e C. Y. Suen. **Methods for Combining Multiple Classifiers and their Applications to Handwriting Recognition**. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(3), pp. 418-435, 1992.