

Marcos Paulo Sanchez

Universidade de Taubaté

Marcio Abud Marcelino

*UNESP, Campus de Guaratinguetá e
Universidade de Taubaté*

Francisco José Grandinetti

*UNESP, Campus de Guaratinguetá e
Universidade de Taubaté*

Correspondência/Contato

UNIVERSIDADE DE TAUBATÉ
Departamento de Engenharia Mecânica

Rua Daniel Danelli, s/n, Jd. Morumbi
Taubaté - SP
CEP 12060-440
Fone (12) 3625-4193

Editores responsáveis

Prof. Dr. Luiz Eduardo Nicolini do P. Nunes
luiz.nunes@unitau.com.br

Profa. Dra. Valesca Alves Correa
valesca.correa@unitau.com.br

ESTUDO DA IDENTIFICAÇÃO BIOMÉTRICA POR IMPRESSÃO DIGITAL

RESUMO

A necessidade um controle rápido e eficiente no registro de frequência do corpo discente de instituições de ensino mostrou uma grande oportunidade de automação com a utilização do acesso biométrico por impressão digital, tornando o controle de frequência eficiente, íntegro e ágil, otimizando o tempo de aula e garantindo a disponibilidade da informação. Este trabalho estudou a utilização da biometria como mecanismo de segurança dentro de sala de aula na identificação de alunos, aplicando o algoritmo linear a partir de um universo selecionado. A busca por uma identificação biométrica deve ser feita registro a registro, o que aumenta e muito o custo da aplicação, desta forma aplicamos o algoritmo linear dentro de um universo selecionado, através de uma chave, diminuindo e muito o custo de busca de um registro. Outro fator é possibilidade de se diminuir a quantidade de pontos equivalentes, uma vez que a comparação é executada dentro de um universo menor. Os resultados se mostraram favoráveis uma vez que a performance de busca dentro do universo total sofre depreciação cada vez que este universo aumenta e a busca linear dentro de um universo selecionado mantém a performance que inicialmente já é superior, uma vez que este universo selecionado mantém uma média linear entre todas as turmas dentro de uma instituição de ensino. Outro fator importante para os resultados na utilização do algoritmo linear é que o comparador de características pode ser otimizado, diminuindo a quantidade de características a serem comparados, uma vez que o universo selecionado para comparação sendo menor, a probabilidade do encontro de impressões digitais com características semelhantes é bastante reduzido.

Palavras-chave: Biometria, Digital, Frequência

ABSTRACT

The abstract should be written in just one paragraph (usually between 100 and 150 words) using an objective and concise discourse with respect to the original article's content and structure. It is relevant condensed peace of information that reflects the article's significant information, i.e., research goals, methodology, tests and results, insights and conclusions. The abstract paragraph must show a direct association with the article keywords. It should not be confused with the introduction and must not contain abbreviations, symbols, formulas, diagrams, footnotes, references to literature or figures. Besides, the description of e personal criticism or points of view is not acceptable.

Keywords: Keyword list. Separated by semicolons. Initial capital letter. From three to five keywords.

1. INTRODUÇÃO

A necessidade um controle rápido e eficiente no registro de frequência do corpo discente de instituições de ensino mostrou uma grande oportunidade de automação com a utilização do acesso biométrico por impressão digital, tornando o controle de frequência eficiente, íntegro e ágil, otimizando o tempo de aula e garantindo a disponibilidade da informação.

O controle de frequência em sala de aula traz dois paradigmas a serem quebrados:

- O docente necessita de tempo para realização da chamada, uma vez que o docente necessita identificar cada discente presente e registrar em seu diário;
- A integridade dos dados perde a confiabilidade com listas de presenças assinadas pelos alunos, uma vez que o professor não será capaz de identificar se os discentes presentes assinaram apenas em seu nome, então o docente ou se despende de boa parte da aula para acompanhar aluno a aluno na assinatura de listas.

A produção das informações físicas geradas, seja em diários ou listas de presenças, deverão ser imputadas no sistema de gerenciamento da instituição e estas tarefas exigem um retrabalho para digitação ou digitalização dos dados coletados, o que decorrerá em intervalos longos de tempo até a apuração total dos dados e disponibilização para consulta ou processamento.

A biometria¹ é uma ciência que possibilita o reconhecimento e identificação por uma análise de características físicas de uma pessoa tais como geometria da mão, impressão digital, íris, retina, reconhecimento facial e voz, entre algumas outras. Sendo estas características consideradas únicas de pessoa para pessoa (DOS SANTOS, 2007) (PINHEIRO, 2008).

Este trabalho sugere a utilização da identificação biométrica por impressão digital na identificação dos discentes em sala de aula para registro de frequência, sendo que tal identificação deva ser realizada de forma que o discente necessite apenas inserir sua digital em um leitor biométrico e o sistema identifique automaticamente sua identidade, registrando então sua presença na aula, não havendo qualquer outro tipo de interação, como, por exemplo, a digitação de seu registro acadêmico. Inicialmente este processo requer um tempo maior de processamento, pois a impressão digital colhida terá que ser comparada, uma a uma, com cada impressão digital armazenada no sistema, uma vez que o formato de armazenamento da informação biométrica não permite a indexação ou busca imediata a partir de filtros ou condições.

2. OBJETIVOS

O objetivo deste trabalho é otimizar o processo de localização da informação biométrica armazenada em um banco de dados, de modo a reduzir o tempo de identificação na comparação entre a informação coletada e as informações armazenadas, utilizando para isso o algoritmo de busca linear dentro de um universo menor de informações biométricas armazenadas, filtrando o universo através das informações de ambiente já registradas no sistema, como o cruzamento do professor que abriu o sistema com o horário de aula registrado, carregando assim apenas os dados biométricos das turmas selecionadas. Outro fator é a regulação do número de coincidências

¹ Biometria – do grego Bios = vida, metron = medida;

necessárias para identificação de um indivíduo uma vez que o universo é menor a probabilidade de coincidências entre impressões, também, é menor.

3. METODOLOGIA

Este estudo visa à análise da forma de identificação digital e o algoritmo de leitura das informações de forma a utilizar as informações ambientais registradas no sistema para diminuição do universo de leitura de informações biométricas aumentando desta forma o tempo de busca por registro no banco de dados.

No sistema de identificação biométrica para controle de frequência de discentes foi utilizado o algoritmo linear, uma vez que se trata de um algoritmo ótimo para localização de dados não indexados. Para otimização do processo, foram utilizadas as variáveis de ambiente como curso, turno, turma e docente, pré-cadastradas no sistema, e obrigatórios para o sistema, para reduzir o universo de informações biométricas a ser consultada, como também a quantidade real de coincidências que devem ser consideradas no comparativo entre a informação biométrica colhida e armazenada.

4. RESULTADOS E DISCUSSÃO

A Biometria é o estudo das características únicas dos indivíduos, utilizado para sua identificação (LIU, 2001). Ela permite a identificação de indivíduos, a partir das características de cada um. Para tanto, ela baseia-se nas características comuns que são distinguíveis entre pessoas diferentes, isto é, características físicas ou comportamentais que são comuns a todos, porém únicas para cada indivíduo.

Foi feita a escolha pela impressão digital por ser o meio de identificação biométrica mais barato e, também, por ser seguro. É a técnica de identificação biométrica mais antiga, e tem sido utilizada com sucesso em inúmeras aplicações, nos dias de hoje temos o exemplo do Registro Geral (RG) ou identidade, utilizadas pelas Secretarias de Segurança Pública em todo o Brasil. A identificação biométrica por impressão digital é definida como segura pelo princípio da unicidade, onde são estimados que a possibilidade de duas pessoas, incluindo gêmeos, terem a mesma impressão digital é menor que um em um bilhão. Essa segurança é aceita pelo meio científico desde 1823 quando o cientista tcheco Jan Evangelista Purkinje durante suas pesquisas sobre glândulas sudoríparas constatou que a pele dos dedos possuía desenhos formados por sulcos e ranhuras únicos para cada indivíduo. (Ashbourn, 2000).

As ranhuras, ou cristas, e os sulcos em cada terminação ou bifurcação de suas linhas formam um desenho geométrico, chamado minúcia. É através das características das minúcias (formato e posição) que a identificação de indivíduos pode ser feita. As minúcias são formadas no feto e acompanham a pessoa por toda vida, sem apresentarem grandes mudanças, desta forma permitem que um perito possa identificar uma pessoa de forma bastante confiável. As minúcias se apresentam em formas bastante particulares, como podemos ver na Figura 1, estas são a base para qualquer sistema de identificação biométrica por impressão digital. Uma imagem de boa qualidade deve encontrar entre 40 e 100 minúcias para identificação.

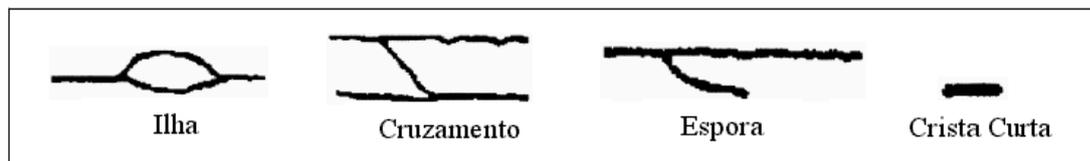


Figura 1. Identificação das minúcias

A premissa para identificação biométrica é a aquisição da imagem da impressão digital do indivíduo, que na figura 2 é chamado de usuário. A partir desta imagem é adquirido o exemplar, deste exemplar são extraídas as minúcias, obtendo-se assim os atributos. (ZHAO e TANG, 2002) A partir destes atributos é gerado um código binário que será armazenado no banco de dados, o que é chamado de *Template*, descrito na Figura 2, por Pinheiro, como Perfil.

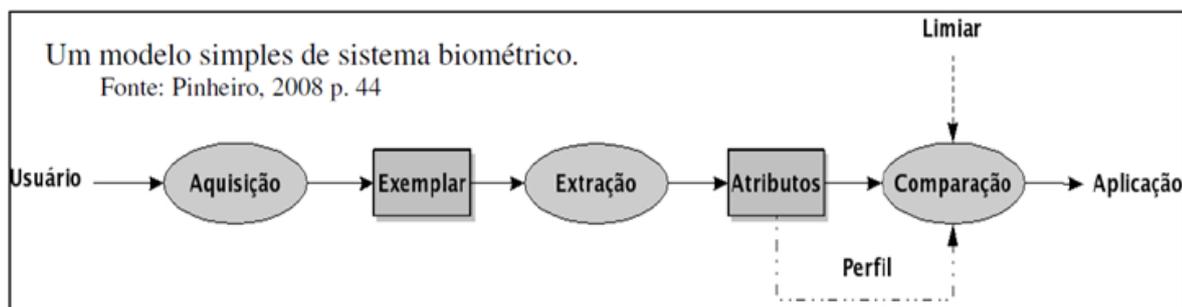


Figura 2. Modelo simples de um sistema biométrico

A comparação dos dados biométricos é feita através de comparações ponto a ponto, minúcia a minúcia, e um a um, ou seja, para encontrar uma digital dentro de um universo armazenado, deve-se comparar o exemplar colhido com cada *template* armazenado (uma a um) e em cada digital comparar cada ponto obtido, verificando o número de coincidências (ponto a ponto). O limiar descrito na Figura 2, indica a quantidade de pontos mínima para o reconhecimento por semelhança, já que na comparação entre o exemplar colhido na hora e a *template* armazenada, nem sempre o número de pontos é igual. A comparação ponto a ponto para cada digital é necessário, por este fato, de que a cada coleta há variações na quantidade e qualidade dos pontos digitalizados, podendo variar posição, iluminação, tipo de sensor, até mesmo a quantidade de gordura pode interferir em uma leitura exata, desta forma a extração exatamente igual dos pontos torna-se inviável devido ao alto custo, tanto de software como de hardware, como também na preparação ambiental para coleta (limpeza e higienização da superfície do dedo, do visor do scanner, iluminação do local, ângulo de leitura etc.). Em uma escola, por exemplo, seria impossível garantir as mesmas condições ambientais a cada leitura, a diversificação de salas, computadores e volume de pessoas é grande, além disso, o tempo para realização da tarefa é extremamente limitado.

Um sistema para identificação biométrica simples é composto por pelo menos dois módulos conforme ilustra a Figura 3, o Módulo de Registro, responsável por gerar a *template* do indivíduo e armazená-la na base de dados e o módulo de Identificação, que é utilizado para coleta e identificação do indivíduo. No caso deste estudo o foco é no módulo de identificação que foi desenvolvido de forma a identificar o discente de forma rápida e segura.

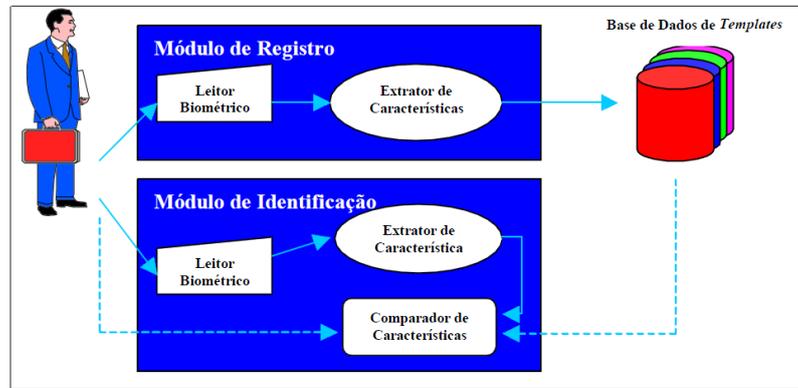


Figura 3. Estágios de um sistema biométrico

Como descrito por Hong (1998, p.4) quando o registro biométrico é cadastrado ele é armazenado em um banco de dados e amarrado a uma chave de identificação do indivíduo. Para o resgate deste registro, pode-se optar por duas formas: 1) o usuário utiliza a chave de identificação, que pode ser o registro acadêmico, CPF, RG etc., para localizar a *template* no banco de dados e então coleta o exemplar para ser comparado, este processo é chamado “um para um”; 2) ou então é feita uma varredura registro a registro no banco de dados até encontrar o *template* semelhante ao exemplar colhido e confirmar a identificação do usuário, este processo é chamado “um para muitos”.

Como visto através da aquisição das características da impressão digital é criado um modelo computacional normalmente denominado *template* biométrico (ver figura 4), este modelo é gerado com propriedades particulares contendo uma síntese de todas as características extraídas com um tamanho adequado de modo a permitir um processo de identificação mais correto e ágil possível.

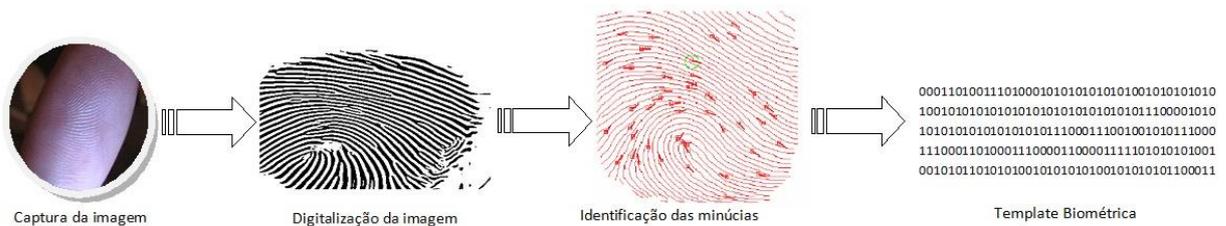


Figura 4. Formação do *Template* Biométrico

A proposta deste estudo foi viabilizar a utilização da identificação biométrica por impressão digital em sala de aula, onde cada sala fosse munida de sensor biométrico conectado a um computador em rede, com a possibilidade de consulta a um servidor de aplicação, onde o sistema biométrico está disponibilizado em um banco de dados responsável por armazenar as *templates* dos discentes (Figura 5).

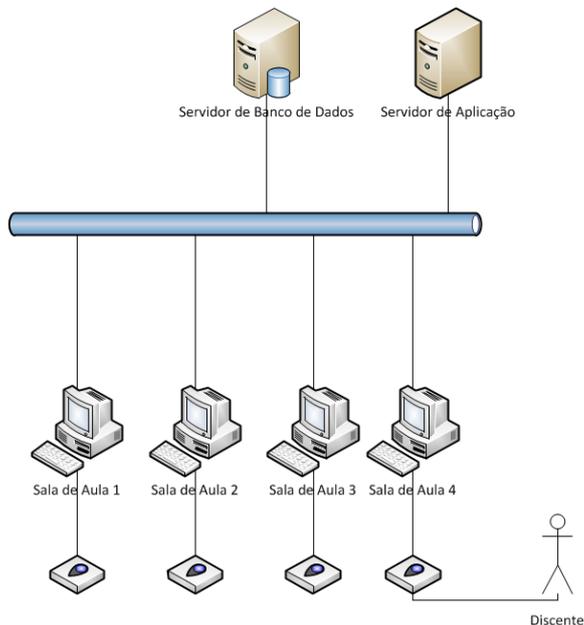


Figura 5. Arquitetura Física Proposta

Os problemas encontrados em ambientes deste tipo é o número de consultas necessárias para identificar um discente e o número de pesquisas concomitantes necessárias para atender a todas as salas de aula. O ambiente considerado no teste é descrito conforme a Tabela 1:

Tabela 1. Dados do Universo utilizado no trabalho

Número de alunos	Templates existentes
8079	16329

Neste universo a Tabela 2, indica o ambiente de teste utilizado, correspondente ao turno noturno com o maior número de alunos e salas simultâneos:

Tabela 2. Turno com maior utilização do sistema

Turno	Salas de Aula	Média de alunos por sala
Matutino	20	45
Vespertino	70	55
Noturno	64	55

Neste ambiente temos o total de tempo necessário para identificação de todos os alunos, conforme a Equação 1:

$$t_T = t_r \times \sum P \tag{1}$$

Onde:

t_T = tempo total necessário para realização de todas as pesquisas;

t_r = tempo de resposta para cada pesquisa;

$\sum P$ = Total de pesquisas realizadas, que é dado por:

$$\sum Alunos \times \sum Templates \tag{2}$$

Considerando os cálculos estabelecidos, os números já mostrados nas tabelas anteriores e definindo $t_r=1$ ms, podemos calcular o tempo necessário para pesquisas em vários ambientes, conforme a Tabela 3:

Tabela 3. Descrição do ambiente inicial de estudo

Ambiente	$\sum Alunos$	$\sum P$	t_T
Em uma sala de aula	60	979.740	16,5 min
Turno matutino	900	14.696.100	244,94 min
Turno vespertino	3850	62.866.650	1.047,78 min
Turno noturno	3520	57.478.080	957,97 min

Como pode ser visto na tabela 3 a melhor das hipóteses apresentada seria a coleta de biometria de uma sala de aula por vez, mesmo assim seriam necessárias 979.740 pesquisas por sala, o que demandaria 16,5 min por sala. Esta situação seria degradada a cada discente acrescentado ao sistema, uma vez que o universo seria aumentado.

A proposta do trabalho visa a criação de um universo de pesquisa único por sala de aula, estabelecendo então uma média de pesquisas, onde independente do universo total de *templates* armazenado, o tempo de busca se manteria dentro de uma média, podendo até estabelecermos um tempo médio de pesquisa por aluno. A figura 5 apresenta a arquitetura proposta, onde é extraída uma base de dados de *template* do banco de dados original e armazenada temporariamente no computador da sala de aula. Esta extração pode ser feita no momento em que o docente aciona o sistema biométrico, a partir daí o sistema verifica as atribuições do docente a partir do horário de aula, e carrega as possíveis *templates* que serão pesquisadas.

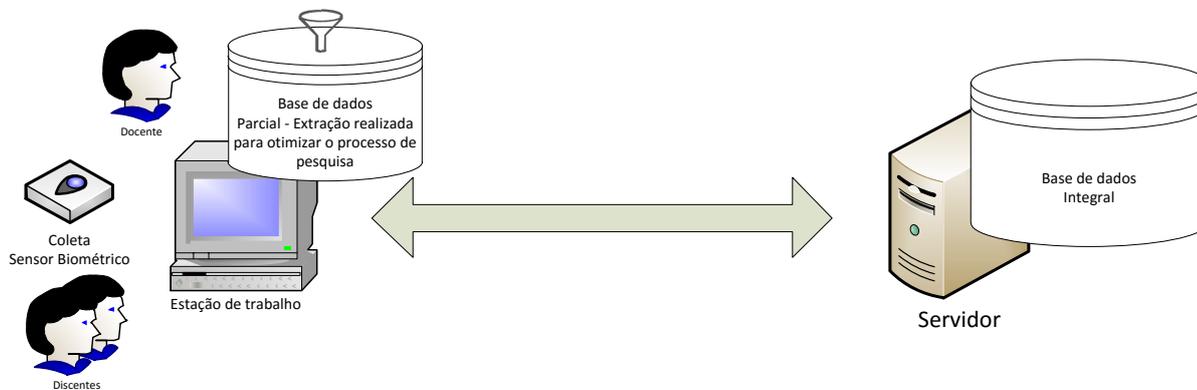


Figura 6. Diminuição do ambiente de pesquisa para otimização do processo de pesquisa

Se fossem consideradas apenas estas modificações, o ganho já seria significativo, conforme apresentado na Tabela 4:

Tabela 4. Ganho

Ambiente	$\sum Alunos$	$\sum P$	t_T
Em uma sala de aula	60	3600	0,05 min

Com a redução do universo de pesquisa a otimização do algoritmo linear, para o ambiente proposto é o fluxograma apresentado na Figura 7:

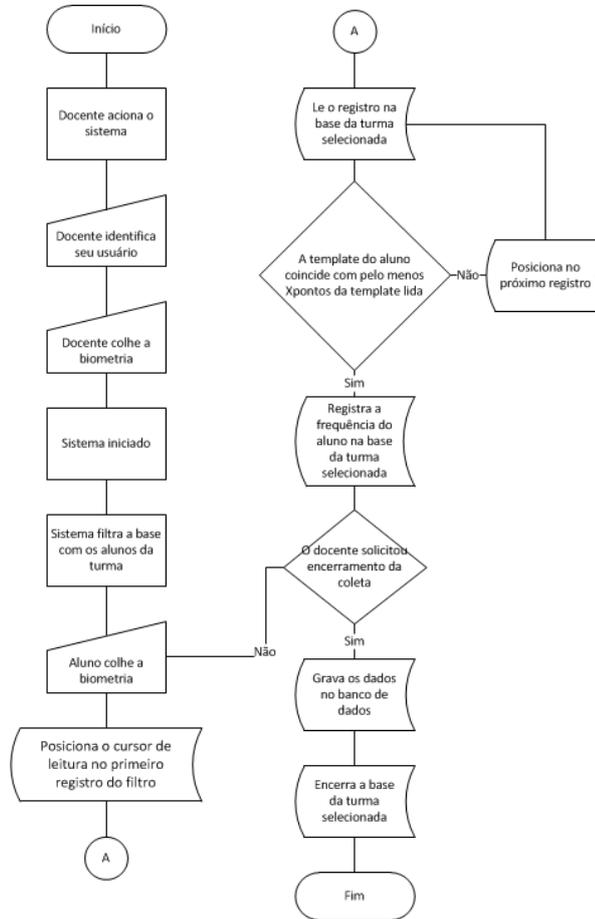


Figura 7. Algoritmo proposto

Outro fator estudado foi o número de coincidências necessárias para afirmar o reconhecimento da identidade de um indivíduo. No processo de comparação ou *matching* a aplicação gera uma pontuação ou *score* indicando quantos pontos foram reconhecidos por semelhança, para definir se houve identificação ou não da impressão digital colhida a aplicação deve ter como parâmetro um limiar a ser considerado, ou seja, o número mínimo de pontos a serem identificados. O aumento ou identificação destes pontos pode influenciar diretamente no tempo de resposta da aplicação. Se considerarmos o número do limiar como número de pontos a serem considerados, ou então fixarmos um teto de números de pontos a serem considerados, pode aperfeiçoar o processo de reconhecimento, uma vez que a comparação de pontos não será mais necessária uma vez que já temos a identificação do indivíduo.

5. CONCLUSÕES

A aplicação do algoritmo linear em universos diminuídos de *templates* permitiu que fosse estipulado um tempo padrão de resposta para o sistema, uma vez que o número de pesquisas ao banco sempre será conhecido, conforme o universo estipulado, sala de aula com até 60 alunos. É importante considerar que mesmo que o universo dobre, o tempo aumentado será relativamente curto, comparado ao de pesquisa ao universo total. Este processo permitiu a manutenção da estabilidade do sistema e a distribuição do processamento do sistema biométrico entre várias estações da instituição, aumentando a performance do sistema e diminuindo a necessidade de potência no servidor de banco de dados e de aplicação, uma vez que uma simples estação pode executar o processo de comparação (*matching*) das *templates* com performance superior ao processamento centralizado.

Outro fator de destaque é a definição de um número de pontos limiar e de teto para efetuar a comparação de pontos da *template* diminuindo o custo de pesquisa e otimizando o processo de reconhecimento biométrico utilizado.

A utilização da biometria permitiu que cada discente pudesse ser identificado de forma única e íntegra, uma vez que a fraude deste tipo de informação traz maiores dificuldades do que assinar a lista ou registrar a presença em um diário. A identificação eletrônica e automática permitiu a disponibilização da informação de forma *on-line* eliminando o trabalho da secretaria de digitação das informações contidas no diário, uma vez que o sistema se auto alimenta. A qualidade das informações coletadas, bem como a confiabilidade da informação disponibilizada teve significativo aumento, uma vez que foi eliminada a possibilidade de erro de digitação.

REFERÊNCIAS

- PINHEIRO, José Maurício. Biometria nos sistemas computacionais, Editora Ciência Moderna, Rio de Janeiro, 2008.
- ASHBOURN, Julian. Biometrics: Advanced Identity Guide The Complete Guide. London: Springer-Verlag, 2000.
- F. ZHAO and X. TANG, "Preprocessing for skeleton-based fingerprint minutiae extraction", CISST Internation Conference, pp. 742-745, 2002.
- JARDINI, Evandro. "Metodologia para Indexação e Busca de Impressões Digitais através do uso de Função de Distância Métrica", Fundação Educacional de Fernandópolis.
- FONTANA, Danilo Rodrigues. "Sistema de Autenticação/Identificação pessoal biométrica através da palma da mão com o auxílio de redes neurais artificiais", Instituto Tecnológico de Aeronáutica, 2009.
- LIU, S.; Silverman, M. "A practical guide to biometric security technology", IT Pro, p. 27-32, 2001.