

A Autenticação de Manuscritos Aplicada à Análise Forense de Documentos

EDSON J. R. JUSTINO¹

FLÁVIO BORTOLOZZI¹

ROBERT SABOURIN²

¹ PUCPR - Pontifícia Universidade Católica do Paraná, Rua Imaculada Conceição, 1155, Curitiba, PR, Brazil - {justino, fborto}@ppgia.pucpr.br

² ÉTS - École de Technologie Supérieure, 1100, rue Notre-Dame Ouest, Montréal, Québec, Canada - sabourin@gpa.etsmtl.ca

Resumo

A análise de documentos questionados oferece um largo espectro de atividades competentes para profissionais ligados às ciências forenses, especialmente criadas para atender às questões judiciais (criminal e cível), relacionadas às fraudes em documentos. No que tange a identificação da autoria de manuscritos, inexistente uma norma padrão nacional ou internacional para a definição de similaridade. As interações com documentos questionados são em geral pessoais, subjetivas e limitadas por algumas regras locais ou culturais. Assim sendo, existe atualmente um direcionamento da comunidade jurídica em busca da definição de padrões [1,2], em bases científicas, para os procedimentos automáticos ou semi-automáticos da análise pericial de documentos questionados.

O objetivo desse artigo é apresentar um estudo das técnicas convencionais utilizadas na análise de documentos questionados relacionados com a autenticação de manuscritos na língua portuguesa. Em seguida, mostrar as principais abordagens utilizadas no desenvolvimento de propostas para os processos de automatização da verificação de manuscritos e por fim apresentar os avanços das pesquisas nessa área na PUCPR.

1. Introdução

A autenticação de manuscritos trata-se de uma subárea da análise de documentos questionados, cujo objetivo é a comprovação ou a dissociação da autoria de um documento manuscrito questionado juridicamente. A análise pericial de documentos manuscritos, como também é conhecida, envolve um conjunto de procedimentos não normalizados, cujo objetivo é permitir a análise e identificação segura da autoria do documento

questionado. A falta de uma norma padrão permite que a análise pericial acabe sendo influenciada pela subjetividade do perito, permitindo que diferentes abordagens executadas sobre o mesmo documento, levem a produção de diferentes laudos.

2. As Classes de Características para Grupos de Escritores

As semelhanças de grafia apresentadas por indivíduos ou grupos de indivíduos que foram ensinados através do mesmo sistema de escrita ou em sistemas semelhantes são conhecidas como classe de características. Estas características podem ser bastante diferentes em relação a outros grupos, mas consistentes dentro do mesmo. Um bom exemplo de classe de características encontra-se no idioma usado pelo escritor questionado.

As classes de características podem servir para reduzir a procura quando se compara um escritor questionado com os padrões de vários escritores diferentes. Contudo, a simples presença de uma classe de características pode identificar um grupo de escritores, mas não um indivíduo. A falta a uma ou mais classes específicas, pode auxiliar na dissociação do escritor do material questionado. No entanto, não proverá uma base formal para uma identificação definitiva.

3. As Características Individuais do Escritor

Acreditar que a escrita de cada autor será sempre distinta dos demais pode não ser totalmente verdade, em alguns podem existir similaridades tanto nas letras quanto nas palavras escritas por diferentes autores, semelhante a um código numérico que pode ser


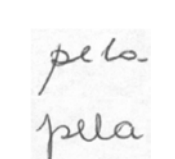




encontrado em diversas caixas fortes, em uma combinação de segurança. No entanto, um código relativamente grande, com uma grande quantidade de números, pode gerar uma combinação sem igual. Portanto, pode existir para cada escritor, uma combinação de características individuais e uma frequência de ocorrências, que fazem com que a escrita de um indivíduo seja distinta dos demais. As diferenciações observáveis na seqüencialidade das letras de uma palavra são características que tendem a conduzir à individualização e como tal, é observada de perto pelo perito durante uma comparação de letras. A seguir são apresentadas algumas dessas características individuais (confira na Tabela 1):


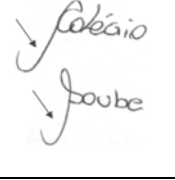


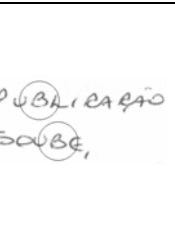

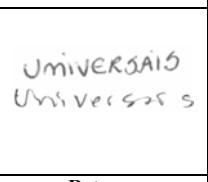
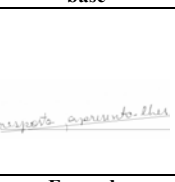
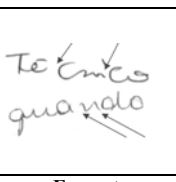



- **Nível de habilidade:** pode ser definido como sendo uma avaliação de beleza aplicada na formação da letra;
- **Inclinação axial:** é o ângulo de inclinação da escrita, em relação ao eixo vertical de um sistema de eixos cartesiano, onde o eixo horizontal é representado por uma linha de base imaginária;
- **Forma caligráfica:** a representação pictórica da escrita;
- **Movimento:** a direção do movimento dos instrumentos de escritura (lápiz ou caneta) pode ser determinada, através da observação das variabilidades na densidade da tinta ou traço do lápis;
- **Proporções:** refere-se geralmente à simetria das letras individualmente;
- **Relações de altura:** são comparação ou correlação da altura de uma letra ou segmento de letra em relação à outra letra, normalmente dentro da mesma palavra;
- **Mínimos gráficos:** são formados pelos pontos finais, vírgulas, os pingos nos “i’s”, acentos (crase, circunflexo, til e agudo) e cedilhas. Essa característica é de grande relevância para os grupos de escritores da língua portuguesa;
- **Corte da letra “t”:** podem contribuir mais significativamente na caracterização do escritor que o pingo da letra “i”;
- **Ascendentes e descendentes:** são laçadas comuns nas letras cursivas. Pode apresentar formas arredondadas ou pontiagudas, simétricas ou assimétricas;
- **Pressão:** representa a variabilidade de largura do traçado e a deposição de material em uma dada região do traço que pode ser tinta ou grafite;
- **Alinhamento em relação à linha de base:** está associada à capacidade do escritor produzir linhas de

textos alinhadas com uma linha guia horizontal fictícia (texto não pautado) ou real (texto pautado);

- **Descontinuidades:** descreve onde o objeto de escrita (caneta ou lápis) ergue-se do papel, normalmente no meio de uma palavra;
- **Velocidade:** é freqüentemente uma característica essencial para a identificação do autor, pois os movimentos rápidos do objeto de escrita são difíceis de ser duplicado por um falsificador;
- **Embelezamento:** localiza-se usualmente no começo de uma letra, mas pode estar presente ao longo do material escrito;
- **Entradas e golpes de saída:** podem ser movimentos habituais, podendo representar características identificadoras de um escritor;
- **Retraço:** considerada uma característica do escritor, enquanto representar um comportamento natural da experiência de escrita;
- **Erros de ortografia e espaçamento:** escrever incorretamente as palavras pode ser um indicativo de uma característica individual do escritor. Existem escritores que interrompem o curso da escrita entre combinações de letras específicas;
- **Formato:** o formato de um documento questionado pode conter adicionalmente uma característica identificadora.

Tabela 1: Exemplos de características individuais do escritor

Nível de habilidade	Inclinação axial	Forma caligráfica
		
Mínimos gráficos	Corte da letra “t”	Ascendentes e descendentes
		

Velocidade	Embelezamento	Entradas e golpes de saída
		
Movimento	Proporções	Relações de altura
		
Pressão	Alinhamento em relação à linha de base	Descontinuidade
		
Retraço	Erros de ortografia e espaçamento	Formato
		

4. Os Modelos para Análise de Manuscritos

Na maioria dos laboratórios forenses, um “exemplar genuíno” ou *standard* é simplesmente um item conhecido com o qual um item desconhecido pode ser comparado. Um exemplar genuíno normalmente possui uma quantidade suficiente de texto escrito para identificar características da individualidade do autor. A origem indubitável da escrita deve ser ligada, nos tribunais, à sua autenticidade. O perito deve confrontar a escrita dos exemplares genuínos com a escrita do documento questionado e com isso produzir um laudo técnico, onde o parecer técnico demonstre sua autenticidade ou discordâncias.

O exemplar original ideal a ser usado para a comparação da escrita manuscrita é aquele obtido sob as mesmas condições com a qual o documento questionado fora produzido. Ele contém as mesmas palavras, números

e símbolos. Foi escrito usando-se aproximadamente o mesmo tempo, e os mesmos tipos de recursos, com o tipo de papel e caneta. O exemplar original deve, portanto, reproduzir suficientemente todas as variabilidades da escrita do autor. Adicionalmente, o mesmo deverá ser produzido sem que o autor conheça o propósito do seu uso. Obviamente, nem todos esses requisitos serão satisfeitos em todos os casos, mas é importante, sempre que possível, que a maioria dessas condições sejam satisfeitas. Se o exemplar original não se desviar demasiadamente da duplicação ideal do documento questionado, pode então o perito utilizar toda a sua habilidade para produzir um laudo definitivo de identificação ou rejeição [3].

Existem dois tipos básicos de exemplares utilizados como modelo, os colhidos e os coletados [4]. Os exemplares colhidos são aqueles documentos de escrita bem simples que foram indiscutivelmente preparados pelo escritor quando o mesmo não tinha razões para pensar que poderiam ser usados em uma demanda judicial. Eles estão, portanto, livres da tentativa de disfarce. A desvantagem do exemplar colhido está na possível dificuldade em encontrar espécimes que reproduzam o formato e texto do documento questionado. Os exemplares colhidos inadequadamente podem conduzir o perito a uma comparação inconsistente. A vantagem do exemplar colhido reside em eliminar a possibilidade do disfarce que freqüentemente supera as possíveis desvantagens.

Os exemplares coletados são aqueles nos quais o indivíduo é intimado a reproduzir um material escrito específico, usualmente através de ditado. Essa classe de exemplares possui a vantagem de conter, aproximadamente, o formato e o conteúdo do documento questionado, produzido de acordo com as orientações do perito. Ele possui, contudo, a desvantagem de o autor conhecer a finalidade do documento, que pode ser usado contra seus interesses.

Existem vários modelos de documentos de coleta, que vêm sendo usados em vários países. Esses modelos, apesar de não duplicarem o conteúdo exato do documento questionado, possuem muitas associações de palavras, letras e símbolos encontradas em cartas comuns. Os modelos se adaptam aos padrões de grafia do idioma usado. Adicionalmente as mesmas apresentam todos os caracteres do alfabeto, maiúsculos e minúsculos, acentuações (característico da linguagem usada), pontuação comum e os números 0 a 9. A maioria das cartas modelo foi confeccionada para grupos de escritores da língua inglesa, como por exemplo, a “Carta de Londres” a “Carta da Classe 16”, entre outras. Portanto, podem vir a desconsiderar características

individuais importantes em um idioma diferente, como por exemplo, os mínimos gráficos encontrados na língua portuguesa.

5. Os Métodos Automáticos de Identificação da Autoria

O objetivo dos métodos automáticos de identificação da autoria é reconhecer o conteúdo da imagem do documento e extrair as informações relativas à autoria ou intenção. Como se trata da autenticidade do documento é natural que as abordagens adotadas, sempre que possível computacionalmente, implemente os procedimentos consolidados da análise grafotécnica pericial. No entanto, como a análise grafotécnica pericial não possui procedimentos rigorosamente padronizados para a definição de similaridades, tais técnicas estão sujeitas às subjetividades e limitações impostas pelas viabilidades computacionais. Por outro lado, é notório que alguns procedimentos periciais estejam consolidados como quesitos mínimos para a obtenção de resultados satisfatórios em uma análise grafotécnica pericial, como visto anteriormente. A busca por esses padrões tem sido observada desde 1998, através da iniciativa do *Federal Bureau of Investigation (FBI)*, através do *Technical Working Group on Forensic Document Examination (TWGDOC)*, que apresentou uma proposta para criar-se uma padronização dos procedimentos usados na identificação da autoria de documentos manuscritos no território americano [2]. Essa iniciativa mostra a necessidade de se definir um conjunto mínimo de quesitos para a validação dos processos automáticos e semi-automáticos associados ao problema.

6. A Verificação Automática de Manuscritos

Os dois principais objetivos das pesquisas na área da análise automática da escrita manuscrita encontram-se primeiro no estabelecimento de normas, com bases científicas, para determinar a individualidade da escrita e em segundo propor soluções automáticas ou semi-automáticas que auxiliem na produção de laudos juridicamente aceitos. Esses dois elementos retiram desse contexto, a subjetividade aplicada pelos peritos durante o processo de análise.

Como visto anteriormente, o principal objetivo da autenticação de manuscritos é constatar a individualidade do escritor, ou a autoria. A autoria da escrita está diretamente relacionada com a intenção do mesmo. Isto é, a análise pericial pode ter como enfoque a identificação de uma fraude ou simplesmente validar a

autoria do documento. Essas diferentes abordagens estão associadas às circunstâncias na qual o documento foi gerado.

No primeiro caso, o processo é mais complexo, pois envolve não somente a certificação da autoria, mas uma possível presença de fraude, como disfarces ou falsificações. Esses tipos de casos são os mais comuns no cotidiano de um perito grafotécnico. Na Figura 1 é possível observar detalhes de um processo de análise pericial, envolvendo recibos de compra. Nesse tipo de análise não se pode descartar a hipótese do disfarce da escrita, o que eleva a complexidade dos processos de verificação.

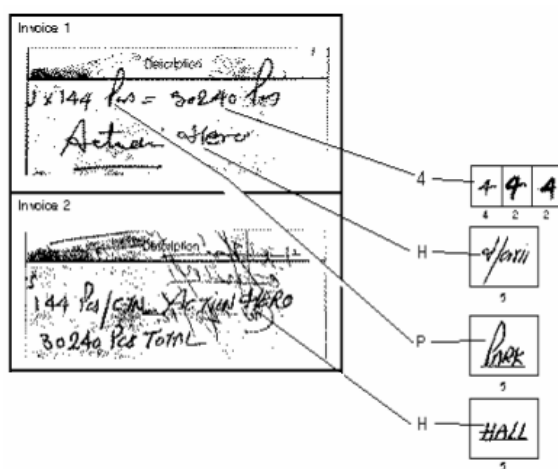


Figura 1: Exemplos de um processo de análise pericial de documentos [5].

No segundo caso, o processo de comparação define o grau de similaridade ou autenticidade de um manuscrito, levando em consideração a escrita natural do autor, isto é, sem a presença de disfarces [6]. Esse tipo de análise pode ser usado, por exemplo, para conferir autenticidade a um documento histórico, confira na Figura 2.

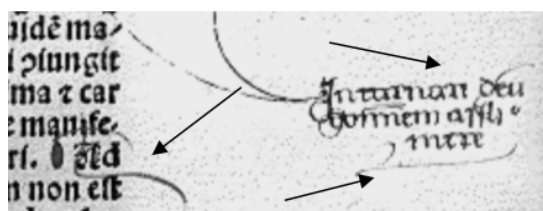
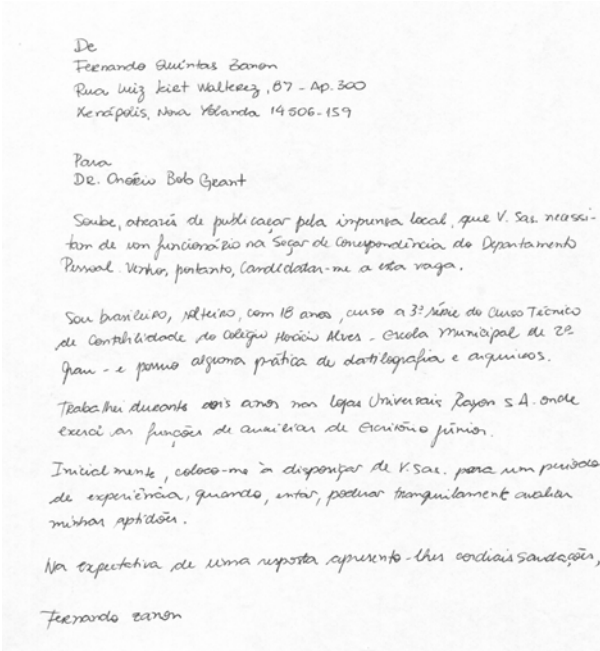


Figura 2: Página do Livro no. 3 da Suma Teológica [Latin Eclesiástico] “St. Thomas Aquinas” de 1436, com anotações na margem direita, usada pelo autor para impor velocidade no traçado [7].

7. As Bases de Dados de Manuscritos

Com o objetivo de desenvolver ferramentas de auxílio no processo de análise pericial, alguns modelos de carta foram criados posteriormente aos modelos tradicionais, descritos anteriormente. Esses modelos estão sendo usados atualmente, para a coleta e formação de bases de dados que permitam a validação dos procedimentos computacionais. O CEDAR (*Center of Excellence in Document Analysis and Recognition*) apresenta um modelo padrão usado na coleta de exemplares de 1000 autores em todo território americano [6].

Um trabalho similar vem sendo executado na PUCPR (Pontifícia Universidade Católica do Paraná) [1], e encontra-se atualmente com 200 autores, confira na Figura 3. Cada autor produz 3 exemplares do documento. O mesmo é composto por todas as letras do alfabeto (maiúsculas e minúsculas), números de 0 a 9, acentos (crase, circunflexo, til e agudo) e cedilhas. A principal diferença entre os dois modelos encontra-se no idioma. A Carta PUCPR apresenta as características da escrita da língua portuguesa, que não estão presentes na carta CEDAR.



De
Fernando Quintas Zanon
Rua Luiz Kirt Walterez, 87 - Ap. 300
Xenópolis, Nova Yolanda 14506-159

Para
Dr. Onório Bob Grant

Soube, através de publicação pela imprensa local, que V. Sas. necessitam de um funcionário na Seção de Correspondência do Departamento Pessoal Venho, portanto, candidatar-me a esta vaga.

Sou brasileiro, solteiro, com 18 anos, curso a 3ª série do Curso Técnico de Contabilidade do Colégio Horácio Alves - Escola Municipal de 2º Grau - e possuo alguma prática de datilografia e arquivos.

Trabalhei durante dois anos nas Lojas Universais Rayon S.A. onde exerci as funções de auxiliar de escritório júnior.

Inicialmente, coloco-me à disposição de V. Sas. para um período de experiência, quando, então, poderão tranquilamente avaliar minhas aptidões.

Na expectativa de uma resposta apresento-lhes cordiais saudações,

Fernando Zanon

(b)

Figura 3: Exemplo de um manuscrito da carta PUCPR, (a) modelo e (b) manuscrito.

8. As Abordagens na Análise Automática de Manuscritos

Existem inúmeras abordagens utilizadas na identificação da autoria da escrita natural [5, 8-10]. Duas se destacam pela forma de tratar o problema da identificação da autoria. A primeira trata da verificação da autoria com base em duas amostras distintas que permitam determinar se pertence ao mesmo autor. Nesse caso, não se avalia a presença de falsificações e nem a possibilidade de disfarce, somente a escrita natural. No segundo caso, uma amostra de manuscrito de um autor desconhecido é comparada com uma base de manuscritos de autores conhecidos, com o objetivo de determinar a autoria. Cada abordagem adota diferentes critérios em cada etapa do processo, segmentação, extração de características, processo de aprendizado e verificação. A Figura 4 mostra um diagrama em blocos de um sistema genérico de verificação de manuscritos.

De
Fernando Quintas Zanon
Rua Luiz Kirt Walterez, 87 - Ap. 300
Xenópolis, Nova Yolanda 14506-159

Para
Dr. Onório Bob Grant

Soube, através de publicação pela imprensa local, que V. Sas. necessitam de um funcionário na Seção de Correspondência do Departamento Pessoal. Venho, portanto, candidatar-me a esta vaga.

Sou brasileiro, solteiro, com 18 anos, curso a 3ª série do Curso Técnico de Contabilidade do Colégio Horácio Alves - Escola Municipal de 2º Grau - e possuo alguma prática de datilografia e arquivos.

Trabalhei durante dois anos nas Lojas Universais Rayon S.A. onde exerci as funções de Auxiliar de Escritório Júnior.

Inicialmente, coloco-me à disposição de V. Sas. para um período de experiência, quando, então, poderão tranquilamente avaliar minhas aptidões.

Na expectativa de uma resposta apresento-lhes cordiais saudações,

Fernando Zanon

(a)

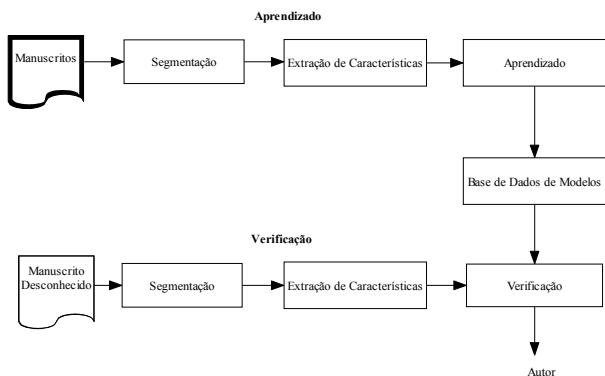


Figura 4: Diagrama em blocos de um sistema genérico para a análise automática de manuscritos.

No aprendizado a imagem do documento modelo passa inicialmente por uma etapa de segmentação, que permite a separação das palavras em cada linha de texto e em seguida, das letras em cada palavra (ver Figura 5). Na etapa de extração de características, cada letra é devidamente rotulada e um conjunto de primitivas é então extraído, com base nas características individuais de cada autor (ver Figura 6). Na etapa de aprendizado é gerado um modelo que descreve as características de um autor, que será posteriormente utilizado na verificação.

Na verificação, a imagem do documento questionado é submetida ao mesmo processo de segmentação e extração de primitivas. Na etapa final, o modelo gerado no aprendizado é então utilizado na confrontação com o modelo questionado.

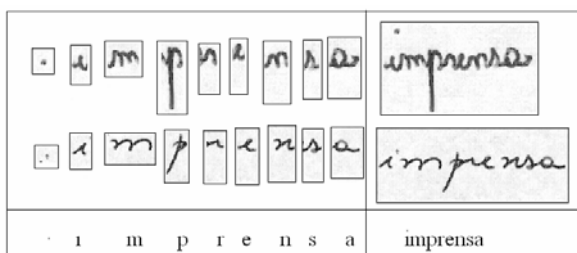


Figura 5: Exemplo de segmentação de palavras em letras.

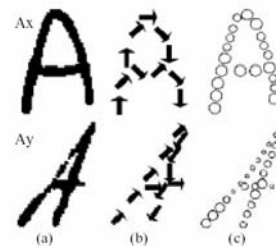


Figura 6: Exemplos de primitivas extraídas dos caracteres: (a) imagem original, (b) característica direcionamento dos segmentos e (c) espessura do segmento [6].

9. Conclusão

O objetivo desse artigo foi apresentar um estudo sobre a autenticação de manuscritos associados à análise forense de documentos. Mostrou-se também, o interesse da área jurídica em se padronizar os procedimentos periciais. Nesse contexto, discutiram-se os princípios básicos que norteiam a análise pericial de documentos manuscritos, a subjetividade associada a esse processo, bem como a complexidade de se desenvolver métodos computacionais para a automatização do mesmo.

Num segundo momento, apresentaram-se os avanços obtidos com as pesquisas na área que visam desenvolver métodos automáticos e semi-automáticos para auxiliar no processo de análise pericial. No entanto, ainda existem problemas em aberto, principalmente no que tange à segmentação automática de palavras, pois os métodos propostos ainda não apresentam resultados adequados ao contexto da autenticidade de manuscritos. Um outro problema reside na dificuldade de se implementar os conceitos da análise pericial em um modelo computacional, tanto em termos de primitivas quanto no processo de tomada de decisão.

Por fim, foram apresentados os avanços das pesquisas nessa área na PUCPR, iniciando-se com a formação da base de dados de cartas forenses, seguido pelo início das pesquisas no desenvolvimento de métodos de segmentação de palavras e letras, e na proposição de métodos que associe as características descritas pela análise grafotécnica pericial e os processos computacionais de automatização.

10. Bibliografia

- [1] E. J. R. JUSTINO, A Análise de Documentos Questionados, Relatório Técnico, Pontifícia Universidade Católica do Paraná, 2002, 74p.
- [2] JEREMY TRAVIS, Forensic Document Examination Validation Studies, Solicitation of National Institute of Justice, Office of Justice Programs, Washington, DC, USA, 1998.
- [3] EMILY J. WILL, Questioned Document Examination Page, www.qdewill.com
- [4] F. HARLEY NORWICH, Norwich Document Laboratory, 17026 Hamlin Boulevard, West Palm Beach, Florida 33470, www.questioneddocuments.com
- [5] JAMES E. DOYLE, State of Wisconsin Department of Justice, www.doj.state.wi.us/index.asp
- [6] SUNG-HYUK CHA, Use of Distance Measures in Handwriting Analysis, PhD. Theses, Faculty of the Graduate School of the State University of New York at Buffalo, April, 2001, 208p.
- [7] HANDWRITING-L GROUP, The Handwriting Analysts Group, www.handwriting.org
- [8] H. E. S. SAID, K. D. BAKER and T. N. TAN, Personal Identification Based on Handwriting, Proc. of ICPR'00, Volume 2, 2000, Barcelona, Spain, p.2797-2800.
- [9] JEAN PIERRE CRETTEZ, A Set of Handwriting Families: style Recognition, Proc. of ICDAR'95, Montréal, Canada, 1995, pp. 489-494.
- [10] T. TAN and Y. ZHU, A Global Texture Analysis for Automatic Font Recognition and Writer Identification in application, Chinese Patent Application No.99105851.8, 1999.