# Reconhecimento de palavras manuscritas referente ao mês do campo da data de cheques bancários brasileiros baseado nos HMMs

Marisa Morita, Edouard Lethelier, Abdenaim El Yacoubi,
Flávio Bortolozzi
Pontifícia Universidade Católica do Paraná
Programa de Pós-Graduação em Informática Aplicada
Laboratório de Análise e Reconhecimento de Documentos
Rua Imaculada Conceição 1155, Prado Velho, 80215-901
Curitiba-Pr, Brazil
{marisa,edouard,yacoubi,fborto}@ppgia.pucpr.br

Robert Sabourin
Laboratoire d'Imagerie, de Vision et d'Intelligence Artificielle
Ecole de technologie supérieure, Département
de génie de la production automatisée
1100, rue Notre Dame Quest, Montréal, H3C IK3, Canada
sabourin@gpa.etsmtl.ca

**Resumo**

Este artigo descreve uma primeira etapa de um sistema em desenvolvimento para o processamento de informações manuscritas de datas no contexto de cheques bancários brasileiros. Esse sistema considera vários tipos de escrita e um contexto *omni*-escritor. Até o presente momento, nós desenvolvemos duas abordagens, global e analítica, para o reconhecimento de palavras isoladas manuscritas com um léxico limitado, referente ao mês do campo da data. Essas abordagens são baseadas na segmentação explícita e no uso dos Modelos de *Markov* Escondidos. Finalmente, nós combinamos ambas as abordagens para poder aproveitar as vantagens de cada uma. Os resultados mostrados nesse artigo são preliminares e os consideramos satisfatórios, dado uma base de dados não muito significativa, o conjunto de primitivas pouco discriminativo, entre outros fatores. No final deste artigo mostraremos algumas perspectivas quanto aos nossos trabalhos futuros.

Palavras-Chave: Inteligência Artificial, Reconhecimento de Padrões, Algoritmos.

# 1 Introdução

Várias pesquisas têm sido realizadas no reconhecimento automático de manuscritos *off-line* para tornarem esses sistemas mais próximos ao comportamento humano. Apesar dos computadores estarem cada vez mais poderosos devido aos grandes avanços tecnológicos nesses últimos tempos, eles ainda são bastante limitados para executarem tarefas como a visão humana. Isso explica o motivo pelo qual as pesquisas nessa área têm sido bastante intensas. Um dos maiores problemas encontrados reside na grande variedade encontrada no estilo da escrita e outros fatores tais como o tamanho do léxico e o número de escritores. Além disso, o reconhecimento *off-line* é mais complexo que o *on-line* pois a imagem é adquirida após a escrita com ruídos, e não pode fazer uso da informação temporal obtida no momento da escrita, como a seqüência ou a velocidade da escrita, que são informações preciosas para um sistema de reconhecimento.

Para diminuir a complexidade do problema, muitos pesquisadores utilizam a informação contextual em seus sistemas para construírem estratégias direcionadas pelo léxico. Em aplicações cujo léxico é pequeno, como o processamento do valor por extenso de cheques bancários podem ser empregadas a abordagem global. Nessa abordagem um modelo para cada classe de palavra é construído separadamente [5] [2]. Entretanto, em aplicações com um grande léxico torna-se inviável treinar um modelo para cada classe de palavra, como é o caso do processamento de endereços postais, devido ao grande número de nomes de cidades e ruas. Em aplicações como essa geralmente é empregada a abordagem analítica, onde um modelo para cada classe de caracter é construído, sendo os modelos das palavras fornecidas pelo léxico construídos pela concatenação de modelos apropriados de caracteres [4] [5] [6] [9].

Na abordagem analítica as palavras são segmentadas em entidades ou *graphemes* que podem representar caracteres ou pseudo-caracteres. Essa estratégia de segmentação é a mais utilizada, devido a grande variabilidade da escrita manuscrita. Nesse caso, os pontos de segmentação definitivos são determinados na fase de reconhecimento pela combinação dos *graphemes*. Apesar da abordagem global não requerer nenhum tipo de segmentação, a abordagem analítica ainda possui vantagens em relação à abordagem global. Uma das grandes vantagens da abordagem analítica reside na sua portabilidade em termos de aplicação, pois para cada nova aplicação não é necessário treinar o novo conjunto de palavras do léxico correspondente. Além disso, é mais eficiente treinar um pequeno conjunto de classes de caracteres do que palavras dependendo do tamanho do léxico.

Os Modelos de *Markov* Escondidos (HMMs) têm sido aplicados com sucesso no reconhecimento da fala. Recentemente, eles têm sido utilizados para o reconhecimento de palavras manuscritas [4] [5] [6] [9] [1]. Esses trabalhos se diferem na arquitetura dos HMMs utilizados e no conjunto de primitivas empregado para representar as imagens de palavras como seqüências de observações. Em particular, algumas abordagens segmentam de maneira explícita as palavras em *graphemes* que podem representar caracteres ou pseudo-caracteres [5] [9] [1]. A segmentação explícita pode ser empregada na abordagem global para determinar o comprimento de uma palavra, de maneira que essa informação mais as primitivas consideradas auxiliem na discriminação das classes de palavras contidas no léxico. Outras abordagens efetuam essa segmentação de maneira implícita, durante o reconhecimento através dos HMMs [4] [6].

O escopo de nosso trabalho consiste no reconhecimento *off-line* de informações manuscritas de datas no contexto de cheques bancários brasileiros, levando em consideração vários tipos de escrita e um contexto *omni*-escritor. A data é composta por palavras (cidade, mês, separadores)

e dígitos (dia e ano), o que torna o nosso sistema de reconhecimento ainda mais complexo. Nós mostraremos nesse artigo os primeiros resultados de nosso trabalho dedicado ao reconhecimento de palavras manuscritas isoladas referentes ao mês. A particularidade do problema do reconhecimento do mês é que apesar do léxico ser pequeno (12 classes), existem grupos de palavras no mesmo que são discriminados entre si somente por uma pequena parte da palavra. Por exemplo, as partes discriminantes das classes de palavras "Setembro" e "Novembro" correspondem a "Set" e "Nov", devido a presença da parte comum "embro" no final das mesmas. Esses tipos de similaridades propiciam uma maior confusão entre as classes de palavras durante o reconhecimento, aumentando a complexidade do mesmo. Nós desenvolvemos ambas as abordagens, global e analítica, as quais são baseadas na segmentação explícita e nos HMMs. Nesse trabalho estamos utilizando nossa base de dados de laboratório que contém imagens binárias com resolução de 300 *dpi* somente com as informações manuscritas de datas no contexto dos cheques bancários brasileiros.

A seção 2 descreve sobre as datas no contexto de cheques bancários brasileiros e a seção 3 descreve sobre os HMMs. A seção 4 detalha o nosso algoritmo de segmentação para fornecer a seqüência de observações a ser utilizada como entrada para os HMMs. A seção 5 mostra as nossas abordagens de reconhecimento, e a seção 6 mostra os resultados obtidos. Finalmente, a seção 7 apresenta nossa conclusão e as perspectivas de nossos trabalhos futuros.

## 2 Aplicação: Datas de Cheques Bancários Brasileiros

Na literatura podemos encontrar vários trabalhos que processam as informações manuscritas dos valores por extenso e numérico de cheques bancários. Entretanto, são pouquíssimos os trabalhos que processam a informação da data. No momento, não encontramos nenhum trabalho que processe essa informação no contexto de cheques bancários brasileiros.

A data é um campo obrigatório no cheque em quase todos os países, sendo que no Brasil a sua não emissão pode ocasionar a devolução do mesmo. A informação da data consiste das seguintes partes, apresentadas abaixo como elas aparecem, da esquerda para a direita:

- nome da cidade (alfabético) opcional;
- **vírgula** (separador);
- dia (numérico);
- **"de"** (separador);
- mês (alfabético);
- **"de"** (separador);
- ano (numérico).

Os três separadores (em negrito) são geralmente impressos nos cheques, bem como as linhas de base. A Figura 1 mostra uma imagem de data extraída de nossa base de dados de laboratório que contém somente as partes manuscritas da data e sem o fundo do cheque. Apesar dos separadores não foram impressos antes da coleta da nossa base de dados, nós observamos várias imagens aonde

os escritores colocaram seus próprios separadores, tais como a vírgula (“,”), (“;”), ponto(“.”) ou (“de”). A Tabela 1 mostra o domínio das três partes obrigatórias do campo da data a serem preenchidas no cheque (dia, mês e ano).

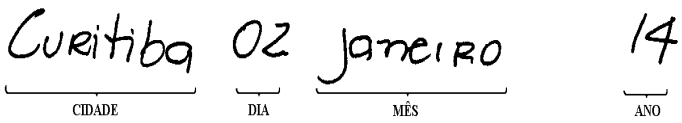


Figura 1: Imagem de data

Tabela 1: Domínio de datas (dia, mês, ano)

| **Dia** | **Mês** | **Ano** |
|---|---|---|
| 01-31 | Janeiro, Fevereiro, Março, Abril, Maio, Junho, Julho, Agosto, Setembro, Outubro, Novembro, Dezembro | 1997-2020 (2 ou 4 dígitos) |

# 3 Modelos de *Markov* Escondidos

Os HMMs [7] tem sido mostrados bem adaptados para caracterizar a variabilidade envolvida em sinais que variam no tempo. A maior vantagem do HMM situa-se na sua natureza probabilística, apropriada para sinais corrompidos por ruídos tal como a fala ou escrita, e na sua fundação teórica devido a existência de um procedimento chamado de algoritmo de *Baum Welch* que iterativamente e automaticamente ajusta os parâmetros do modelo dado um conjunto de treinamento de seqüências de observações. Esse algoritmo garante que o modelo seja convertido ao local máximo da probabilidade de observação do conjunto de treinamento de acordo com o critério de estimação de probabilidade máxima, que depende diretamente dos parâmetros iniciais do HMM.

Em algumas aplicações como a nossa, é mais conveniente que os HMMs emitam as observações entre as transições e não por estado. Consequentemente, é necessário modificar a definição clássica dos HMMs [7], que podem ser encontradas no trabalho de El Yacoubi et al [9]. Nesse artigo estamos considerando os HMMs discretos, dado que as nossas observações são produzidas por um alfabeto.

# 4 Representação de Palavras Isoladas

Uma palavra referente ao mês é segmentada explicitamente em *graphemes*. Assim, para cada *grapheme* da palavra é extraído um conjunto de primitivas que correspondem à um símbolo previamente definido no HMM. Essas primitivas são extraídas com o auxílio das linhas de referência detectadas nas palavras que podem ser obtidas globalmente em toda a imagem ou localmente em cada sub-imagem proveniente de uma segmentação ideal do campo da data ou não. Devido a presença de números no campo da data nós optamos pela detecção local dessas linhas sem buscar

uma segmentação ideal do campo da data. A seqüência ordenada de símbolos será então utilizada como entrada para nossos HMMs.

## 4.1 Segmentação do Campo da Data

Uma imagem de data é segmentada em sub-imagens a partir dos seus espaçamentos significativos (brancos). Esses espaçamentos são detectados a partir do valor médio dos espaçamentos entre dois componentes conectados (CCs) adjacentes situados na região mediana da imagem de data, conforme mostra a Figura 2(a). Essa região é determinada a partir da linha mediana que é calculada através do histograma de transição horizontal. Nesse caso não buscamos uma segmentação ideal em cidade, dia, mês e ano.


Figura 2: (a) Detecção dos espaçamentos significativos contidos na imagem de data e (b) Detecção das linhas de referências em cada sub-imagem

## 4.2 Detecção das Linhas de Referência de Palavras

As linhas de referência detectadas nas palavras correspondem às: linhas de base superior e inferior, linhas superior e inferior, e a linha mediana. Essas linhas delimitam as regiões superior, mediana e inferior de uma palavra, conforme mostra a Figura 2(b), para detectar os ascendentes, o corpo das minúsculas e os descendentes de uma palavra, respectivamente.

A detecção dessas linhas está sendo efetuada para cada sub-imagem da data situada entre dois espaçamentos significativos. A linha de base inferior (superior) é calculada a partir dos pontos de mínimos (máximos) detectados no contorno inferior (superior) dos CCs que foram previamente filtrados através dos mínimos quadrados ponderados. A linha de base inferior (superior) corresponde ao valor médio obtido de todas as ordenadas dos pontos de mínimos (máximos). A linha mediana corresponde a ordenada que passa no meio entre essas duas linhas. As linhas superior e inferior são detectadas a partir dos limites superior e inferior dos CCs, respectivamente.

## 4.3 Segmentação de Palavras em Caracteres ou Pseudo-caracteres

Uma palavra é segmentada em *graphemes* na abordagem analítica. Os pontos de segmentação definitivos são determinados somente na fase de reconhecimento pela combinação dos *graphemes*. No caso da abordagem global esse tipo de segmentação não é requerido, entretanto, ele pode auxiliar na detecção do comprimento da palavra e na detecção da posição de cada símbolo. Por enquanto a palavra referente ao mês está sendo localizada manualmente na imagem de data.

A detecção dos pontos de segmentação (PsS) é baseada nas seguintes hipóteses:

- Alguns caracteres estão naturalmente isolados na palavra (letra "n" da Figura 3);
- Os pontos de mínimos (PsM) do contorno superior correspondem as ligações entre os caracteres.

Os PsS são detectados dentro de uma altura $H$ entre os limites superior e inferior. Esses limites representam 40% da altura da região mediana quando as linhas superior e inferior são maiores que esse valor. Na Figura 3 o valor do limite superior corresponde a linha superior e o valor do limite inferior corresponde a 40% da altura da região mediana. A nova área é então determinada como a região significativa inferior para detectar os descendentes.

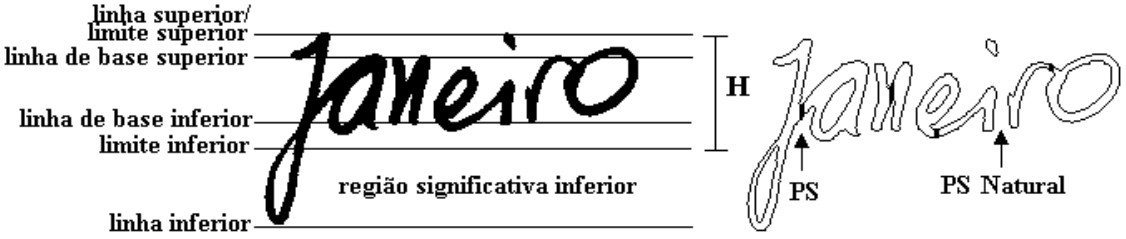


Figura 3: (a) Delimitação da região de segmentação $H$, (b) Tipos de PsS

Para que um PM corresponda a um PS, deve ser possível a sua projeção vertical até o contorno inferior. Essa regra também vale para o PS detectado na vizinhança de um PM nos seguintes casos:

- A projeção vertical de um PM passa por um *loop* (Figura 4(a)) (configuração interna);
- A projeção vertical do PM passa pela tangente do contorno inferior (Figura 4(b)) (configuração tangente);
- O comprimento da projeção vertical do PM até o contorno inferior é maior que um limiar proporcional a espessura do traço $E_t$, que é estimada sobre toda a imagem (Figura 4(c)).

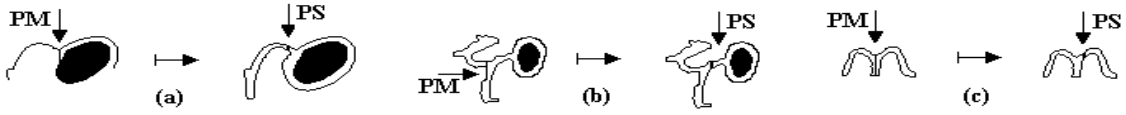


Figura 4: (a) PS situado mais à esquerda do PM com configuração interna, (b) PS situado mais à direita do PM com configuração tangente e (c) PS situado mais à direita do PM com comprimento vertical maior que um determinado limiar

A busca de um PS na vizinhança de um PM inicia no contorno superior à direita do PM. Um PS é detectado no ponto desse contorno dependendo do valor do comprimento da projeção vertical

desse ponto até o contorno inferior. Se não for detectado nenhum PS é considerado aquele ponto com o menor comprimento na vertical. Entretanto, se ainda nenhum PS for detectado à direita do PM devido às configurações interna ou tangente, uma busca é efetuada à esquerda do PM, da mesma maneira que à direita. Os PsS detectados na palavra são filtrados nos seguintes casos:

- A distância no contorno (em índice) entre dois PsS pertencentes ao mesmo CC é menor que um limiar proporcional a $E_t$;
- a distância no contorno (em índice) entre o PS e o início ou final do CC é menor que um limiar proporcional a $E_t$.

O método de segmentação proposto pode produzir a segmentação correta do caracter, a sub-segmentação ou ainda a sobre-segmentação de um caracter em dois ou três *graphemes*.

## 4.4 Extração de Primitivas

Para cada *grapheme* obtido no processo de segmentação é extraído um conjunto de primitivas globais que corresponde aos ascendentes, descendentes e *loops*. Entretanto, um *grapheme* dependendo da escrita pode não ser representado por essas primitivas. Por isso, além dessas primitivas estamos considerando a ausência das mesmas para levar em consideração o comprimento da palavra, e desta forma, aumentar a discriminação das palavras de nosso léxico.

Os ascendentes são detectados a partir dos pontos de máximos no contorno superior de uma palavra e os descendentes são detectados a partir dos pontos de mínimos no contorno inferior da mesma. Nesse caso, estamos discriminando os grandes descendentes (ascendentes) dos pequenos descendentes (ascendentes).

Os descendentes (ascendentes) são localizados na região significativa inferior (superior). A região dos pequenos descendentes (ascendentes) está situada entre o limite inferior (superior) e este limite acrescido (decrescido) de 40% da altura da região mediana, sendo que a parte restante é considerada a região dos grandes descendentes (ascendentes), conforme mostra a Figura 5. Esse limiar empírico foi o melhor entre vários limiares testados na fase de reconhecimento e observados durante a análise de erros no conjunto de validação.

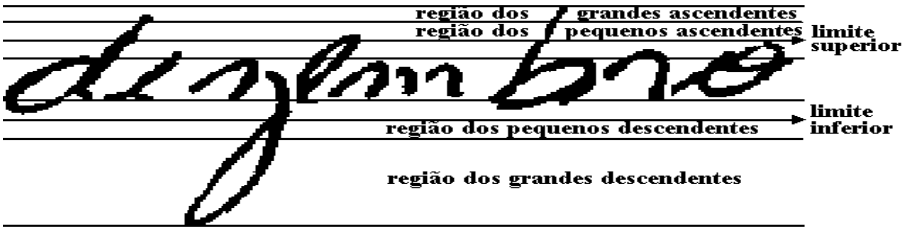


Figura 5: Detecção das regiões dos grandes e pequenos descendentes (ascendentes)

Os *loops* podem ser detectados tanto na região superior, mediana ou inferior. Entretanto, só vai existir um *loop* na região inferior (superior) se existir um descendente (ascendente) correspondente. Do contrário esse *loop* é considerado dentro da região mediana. Somente se o *loop* estiver localizado dentro da região mediana ele é classificado em grande ou pequeno.

Ao contrário das primitivas propostas por El Yacoubi et al em [9], a ordem horizontal de um *loop* na região mediana e o ascendente ou descendente dentro de um mesmo *grapheme* não está sendo considerada para discriminar os caracteres "b" e "d" ou "p" e "q", pois os dois primeiros caracteres aparecem em posições bem diferentes nas palavras de nosso léxico e não há ocorrência dos dois últimos caracteres.

O agrupamento das primitivas globais ou a ausência dessas formam o nosso alfabeto que é composto por 20 símbolos que foram definidos de acordo com a ocorrência de cada primitiva ou o agrupamento das primitivas dentro de cada *grapheme* de uma palavra em todo o conjunto de treinamento, para que haja um número suficiente de símbolos para o aprendizado dos HMMs.

# 5 Reconhecimento de Palavras Isoladas

Quando o conjunto de palavras a ser reconhecido é pequeno e constante, é possível treinar cada palavra do léxico separadamente, como é o caso do reconhecimento de palavras referentes ao mês do campo da data. Entretanto, quando o léxico é pequeno, mas a base de dados não é muita significativa como a nossa, pode ser interessante empregar a abordagem analítica devido ao conjunto de treinamento possuir uma maior freqüência em termos de caracteres do que palavras. Portanto, nós desenvolvemos ambas as abordagens, global e analítica, e em seguida combinamos os resultados dessas abordagens para poder aproveitar as vantagens de cada uma.

## 5.1 Modelos propostos

Em ambas as abordagens empregadas em nosso trabalho adotamos o modelo esquerda-direita ou modelo de *Bakis*, da mesma maneira que uma palavra é escrita. Em nossos HMMs, as observações são emitidas entre as transições. Esse tipo de arquitetura é mais conveniente para as nossas abordagens de modelagem baseada na segmentação explícita.

Na abordagem analítica construímos um modelo de caracter composto por 4 estados baseado na arquitetura proposta por El Yacoubi et al em [9], conforme mostra a Figura 6(a). Essa arquitetura modela a segmentação de um caracter em até 3 *graphemes*. A transição $t_{03}$ modela o caso da sub-segmentação de um caracter, emitindo o símbolo nulo $\emptyset$ ou então modela a segmentação correta de um caracter. A transição $t_{23}$ modela o caso da sobre-segmentação de um caracter em dois *graphemes*, emitindo o símbolo nulo $\emptyset$ ou então modela o caso da sobre-segmentação de um caracter em três *graphemes*. Esse modelo de caminho paralelo absolve melhor a homogeneidade dos *graphemes* produzidos pela segmentação. Nessa abordagem foram desenvolvidos dois modelos para cada classe de caracter, um modelo para o caracter minúsculo e o outro maiúsculo, totalizando 40 HMMs. O número de classes de caracteres considerado é 20, pois nem todos os caracteres estão presentes em nosso léxico.

A arquitetura adotada na abordagem global permite a transição de um estado para os três estados mais próximos, se existirem, conforme mostra a Figura 6(b). Nesse caso, as transições não possuem um relacionamento direto com o comportamento de segmentação ao nível do caracter, tornando o modelo mais flexível para absolver melhor a sobre-segmentação e a interação entre caracteres dentro de uma mesma palavra. Nessa abordagem foram desenvolvidos dois modelos para cada classe de palavra de nosso léxico, um modelo para as palavras minúsculas e o outro para

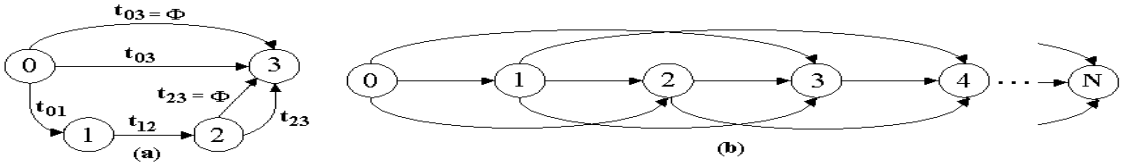


Figura 6: (a) Modelo de caracter, (b) Modelo de palavra com N estados

as palavras maiúsculas, totalizando 24 HMMs. O caso *mixed* não foi considerado devido a nossa base de dados não ser muito significativa. Uma palavra é considerada maiúscula se pelo menos a metade dos caracteres que a compõem são maiúsculos.

## 5.2 Fase de treinamento

O objetivo da fase de treinamento é de estimar os melhores parâmetros dos modelos, dado um conjunto de treinamento de seqüência de observações. Esses parâmetros estão sendo ajustados pelo procedimento de *Baum Welch* com *Cross Validation* com algumas variâncias em relação a definição formal clássica dos mesmos [9]. Este procedimento permite monitorar o desempenho geral durante a aprendizagem dos modelos a partir dos conjuntos de treinamento e de validação.

Devido os modelos na abordagem analítica serem construídos a nível de caracter, os modelos das palavras são formados pela concatenação de modelos apropriados de caracteres, dado um conjunto de treinamento com as palavras rotuladas. Nesse caso, o estado final de um modelo de caracter torna o estado inicial do próximo modelo, e assim sucessivamente, conforme mostra a Figura 7. Ao contrário dos trabalhos descritos em [4] [6], os limites de cada caracter não precisam ser identificados manualmente dentro da palavra.

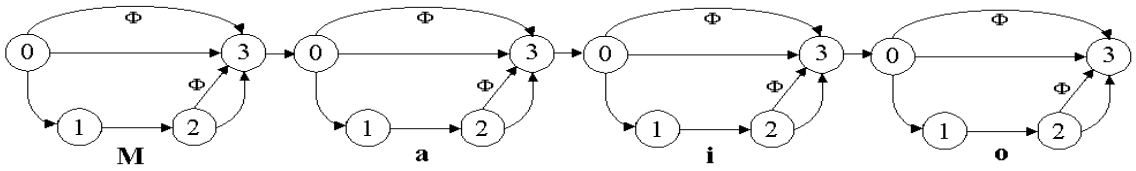


Figura 7: Modelo de treinamento da palavra Maio

Ao contrário do modelo de caracter proposto, que possui um número definido de estados, o modelo global de cada palavra define o número adequado de estados, empregando o procedimento de *Baum Welch* com *Cross Validation* para diferentes números de estados sobre os mesmos conjuntos de treinamento e validação. O número de estados que melhor se adapta a esses conjuntos corresponde ao modelo com a maior probabilidade de validação.

## 5.3 Fase de reconhecimento

Os modelos de reconhecimento das palavras de nosso léxico na abordagem analítica estão sendo criados basicamente da mesma maneira que no treinamento. Entretanto, como a informação sobre

o estilo da escrita (maiúscula, minúscula) de uma palavra desconhecida a ser reconhecida não está disponível, nós adotamos três configurações de palavras: compostas somente por caracteres maiúsculos ou então minúsculos e palavras minúsculas começando com um caracter maiúsculo. Desta maneira, a probabilidade final é obtida pela combinação (soma) das probabilidades produzidas por esses três modelos, considerando a probabilidade a *priori* de cada modelo que corresponde a freqüência de cada classe no conjunto de treinamento. De maneira similar, os modelos globais de palavras maiúsculas e minúsculas são combinados na abordagem global.

# 6 Experimentos de reconhecimento e análises

A base de dados de laboratório contém no total 1.999 imagens de datas, sendo que 1.188 (407) dessas imagens foram utilizadas no conjunto de treinamento (validação). A distribuição de palavras minúsculas é em torno de 80% contra 20% de palavras maiúsculas.

Os experimentos de reconhecimento foram efetuados em um conjunto de testes com 404 imagens de datas. Nós utilizamos o algoritmo de *Viterbi* e o procedimento de *Forward*, mas a Tabela 2 detalha os melhores resultados obtidos com o procedimento de *Forward*. A terceira e a quarta coluna mostram as taxas de reconhecimento obtidas nas abordagens global (G%) e analítica (A%), para cada classe de palavra ("1" para "Janeiro", etc). A última coluna (C%) mostra a combinação (soma) dessas abordagens, e a última linha mostra as taxas médias de reconhecimento.

Tabela 2: Resultados de reconhecimento no conjunto de testes

| **Classe** | $n^0$ **Imagens** | **G%** | **A%** | **C%** |
|---|---|---|---|---|
| **1** | 39 | 82,05 | 92,31 | 94,87 |
| **2** | 32 | 78,13 | 78,13 | 81,25 |
| **3** | 36 | 69,44 | 72,22 | 69,44 |
| **4** | 39 | 87,18 | 89,74 | 87,18 |
| **5** | 38 | 78,95 | 81,58 | 78,95 |
| **6** | 30 | 80,00 | 83,33 | 80,00 |
| **7** | 33 | 78,79 | 60,61 | 84,85 |
| **8** | 28 | 89,29 | 82,14 | 89,29 |
| **9** | 31 | 87,10 | 80,65 | 87,10 |
| **10** | 30 | 86,67 | 93,33 | 93,33 |
| **11** | 34 | 91,18 | 79,41 | 88,24 |
| **12** | 34 | 70,59 | 64,71 | 70,59 |
| **Total** | 404 | 81,43 | 79,95 | 83,66 |

Pode-se verificar na Tabela 2 que a taxa média de reconhecimento foi um pouco menor na abordagem analítica, que possui uma freqüência maior em termos de caracteres no conjunto de treinamento do que de palavras na abordagem global. Essa taxa foi maior na abordagem global provavelmente devido aos seus modelos serem mais flexíveis, pois nesse caso o número adequado de estados de cada modelo é escolhido durante o treinamento e é aquele que vai melhor se adaptar aos

conjuntos de treinamento e validação. Além disso, esses modelos podem absolver melhor a interação entre os caracteres dentro de uma mesma palavra, devido principalmente a grande dificuldade em manter uma regularidade na segmentação do caracter em palavras cuja variabilidade é grande como a nossa.

A particularidade do reconhecimento de palavras manuscritas referente ao mês é que apesar do léxico ser pequeno, existem grupos de palavras no mesmo que são discriminados entre si somente por uma pequena parte da palavra devido a presença de partes comuns, tais como:

- A terminação "eiro" de "Janeiro" e "Fevereiro";
- A terminação "embro" de "Setembro", "Novembro" e "Dezembro";
- Quase todos os caracteres entre "Junho" e "Julho" e entre "Maio" e "Março".

Esses tipos de similaridades propiciam uma maior confusão entre as classes de palavras durante o reconhecimento. As Figuras 8(a) e (b) mostram exemplos de imagens corretamente classificadas e as Figuras 8(c) e (d) mostram alguns contra exemplos.


Figura 8: (a) e (b) Imagens corretamente classificadas, (c) e (d) Erros de classificação

No momento é bastante difícil comparar o nosso trabalho com outros, pois a literatura indica que muitos poucos estudos tem sido efetuados no processamento da informação da data, os únicos trabalhos que conhecemos são os desenvolvidos por Fan et al [3], e uma extensão do mesmo desenvolvidos por Suen et al [8], ambos no contexto de cheques canadenses escritos em francês ou inglês, aonde o mês pode ser grafado numericamente e escrito antes que o dia. Além disso, o foco principal desses trabalhos é a segmentação do campo da data em dia, mês e ano.

# 7 Conclusão e Perspectivas

Este artigo descreve uma primeira etapa de um sistema em desenvolvimento para o processamento *off-line* de informações manuscritas de datas no contexto de cheques bancários brasileiros que considera vários tipos de escrita e um contexto *omni*-escritor. Uma das maiores dificuldades encontradas nesse sistema reside na própria aplicação em si devido o conteúdo da data ser composto de dígitos e palavras. No momento não encontramos nenhum trabalho que processe essa informação no contexto de cheques bancários brasileiros.

Nossos esforços até o presente momento concentraram-se no reconhecimento de palavras manuscritas referente ao mês do campo da data. Nesse caso, como o léxico é pequeno e constante, foi possível treinar um modelo para cada classe de palavra, empregando a abordagem global. Entretanto, como a nossa base de dados não é muito representativa, empregamos a abordagem analítica

para obter uma maior freqüência em termos de caracteres do que palavras no conjunto de treinamento. Desta maneira podemos aproveitar as vantagens de cada uma. Essas abordagens são baseadas na segmentação explícita e no uso dos HMMs.

A taxa média de reconhecimento obtida pela combinação das abordagens global e analítica foi em torno de 83%. Nós consideramos esses resultados preliminares satisfatórios, dado uma base de dados não muito significativa e levando em consideração que o nosso algoritmo de segmentação e o conjunto de primitivas empregado são mais adaptados para as palavras cursivas.

Uma das perspectivas de nossos trabalhos futuros é empregar outros conjuntos de primitivas para melhorar a discriminação entre as classes de palavras. Além disso, desenvolver um modelo de reconhecimento que consiste de dois modelos de caracteres em paralelo, um maiúsculo e o outro minúsculo, para considerar as palavras escrita na forma *mixed* e finalmente efetuar o reconhecimento das outras partes do campo da data como o dia e o ano.

# Referências


[1] M. Y. Chen, A. Kundu, and S. N. Srihari. Variable duration hidden markov model and morphological segmentation for handwritten word recognition. *IEEE Transactions on Image Processing*, 4(12), December 1995.

[2] G. Dimauro, S. Impedovo, G. Pirlo, and A. Salzo. Automatic banckcheck processing: A new engineered system. *International Journal of Pattern Recognition and Artificial Intelligence*, pages 467–503, 1997.

[3] R. Fan, L. Lam, and C. Y. Suen. Processing of date information on cheques. In *Fifth International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pages 207–212, September 1996.

[4] A. M. Gillies. Cursive word recognition using hidden markov models. In *Fifth U.S. Postal Service Advanced Technology Conference*, pages 557–562, 1992.

[5] M. Gilloux, M. Leroux, and J. M. Bertille. Strategies for handwritten words recognition using hidden markov models. pages 299–304, 1993.

[6] M. A. Mohamed and P. Gader. Handwritten word recognition using segmentation-free hidden markov modeling and segmentation-based dynamic programming techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(5), May 1996.

[7] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *IEEE*, 77(2):257–286, February 1989.

[8] C. Y. Suen, O. Xu, and L. Lam. Automatic recognition of handwritten data on cheques - fact or fiction ? *Pattern Recognition Letters*, 20(13):1287–1295, November 1999.

[9] A. El Yacoubi, R. Sabourin, M. Gilloux, and C. Y. Suen. Off-line handwritten word recognition using hidden markov models. In *Knowledge Techniques in Character Recognition*. CRC Press LLC, April 1999.