

CODIFICADORES DE CADEIA EFICIENTES PARA COMPACTAÇÃO DA INFORMAÇÃO DE FORMA EM COMPRESSÃO DE VÍDEO BASEADA EM OBJETOS

Emilio Carlos Acocella e Abraham Alcaim

Resumo - O codificador de cadeia diferencial (CCD) é uma das estratégias mais utilizadas para codificação de forma, sem perdas, em aplicações de compressão de vídeo baseada em objeto. Neste artigo são identificadas e avaliadas diversas modificações de um CCD, que visam o aumento de eficiência dessa estrutura. Nesse sentido, aumentou-se o número de classes de predição espacial em um CCD condicional e buscou-se explorar de forma mais efetiva e integrada as correlações espaciais e temporais entre os pixels de contorno do objeto. As modificações que proporcionaram maior compactação da informação de forma, em relação aos métodos encontrados na literatura, foram incorporadas aos três esquemas propostos neste artigo: o CCD condicional adaptativo, o CCD condicionado ao quadro corrente e o CCD condicionado ao quadro anterior. Em todos os casos foram também considerados os efeitos de cada uma das possíveis escolhas da tabela de códigos de Huffman sobre a taxa de compressão final.

Palavras-chave: Processamento de Imagens, Compressão de Vídeo Baseada em Objetos, Codificação de Forma.

Abstract - In applications of object-based video compression, one of the most efficient strategies for lossless coding of object shapes is the differential chain coder (DCC). In this paper, several modifications of a DCC are identified and analyzed with the purpose of improving its efficiency. The number of spatial prediction classes in a conditional DCC was increased. In addition, we explored in a more effective and integrated form, the spatial and temporal correlations between pixels of the object contour. As compared to the methods available in the literature, the modifications that provided the highest compaction of the shape information were incorporated to the three schemes proposed in this paper: the adaptive conditional DCC, the DCC conditioned to the current frame and the DCC conditioned to the previous frame. In all cases, we have also considered the effects of different choices for the table of the Huffman codes, over the final compression rate.

Keywords: Image Processing, Object-Based Video Compression, Shape Coding.

1. INTRODUÇÃO

Além da codificação da intensidade (textura) e dos parâmetros de movimento, a codificação de forma é um dos pré-requisitos essenciais da representação digital de vídeo baseada em objeto. Dada a importância da forma para a qualidade subjetiva da imagem, e a necessidade dessa informação para a codificação de vídeo baseada em objeto, métodos eficientes para codificação (ou compactação, no caso sem perdas) têm atraído muita atenção nos últimos anos. Esse interesse tem sido especialmente estimulado pelo papel fundamental da representação de forma no padrão MPEG-4 para compressão digital de vídeo [1].

Os codificadores baseados em mapas de bits indicam, para cada pixel, se ele pertence ou não ao objeto. Os baseados em contornos codificam o contorno do objeto. O segundo tende a ser mais eficiente e menos redundante, pois o mapa de bits correspondente a um objeto pode ser obtido do seu contorno pelo preenchimento do seu interior. Isso tem motivado um grande número de trabalhos sobre representação de forma baseada em contornos.

Uma das estruturas mais utilizadas para codificação de forma baseada em contornos sem perdas é o codificador de cadeia. Em particular, o codificador de cadeia diferencial (CCD) [2] tem se mostrado uma abordagem muito atraente e amplamente empregada em aplicações de compressão de vídeo baseada em objeto. Uma sugestão interessante e importante nesse contexto é o CCD condicional (CCDC) [3]. Esse esquema usa a probabilidade da direção diferencial do pixel atual, condicionada à classe a que pertence a direção diferencial do pixel anterior. A partir dessa probabilidade é determinado um valor predito para a direção diferencial e um erro de predição, o qual é codificado por um codificador de entropia.

Neste artigo nós tiramos maior proveito das correlações espaciais entre pixels de contorno e combinamos com estratégias que exploram correlações temporais entre esses pixels. Mais especificamente, utilizamos condicionamentos de segunda ordem e estatísticas do quadro anterior. Com isso, as probabilidades condicionais são estimadas de forma adaptativa, sem a necessidade de transmiti-las. Outro aspecto importante é que diferentes formas de obtenção da tabela de Huffman têm efeitos distintos sobre a eficiência dos codificadores descritos neste artigo. Resultados relacionados a esses efeitos são também apresentados e discutidos.

Na Seção 2 deste artigo é feita uma breve descrição do CCD e de sua versão condicional (CCDC). O restante do artigo trata dos trabalhos que foram desenvolvidos com base na estrutura CCDC. A Seção 3 apresenta as mudanças introduzidas para o aumento do desempenho dos codificadores de

Emilio Carlos Acocella é do Ministério da Defesa, Rio de Janeiro, RJ, Brasil. Abraham Alcaim é do CETUC - PUC, Rio de Janeiro, RJ, Brasil. E-mails: diretor.parque1@uol.com.br, alcaim@cetuc.puc-rio.br. Editor de Área responsável: Ricardo M. Campello de Souza. Artigo submetido em 09/Set/2002, revisado em 09/Jun/2003, aceito em 19/Nov/2003.

cadeia diferencial encontrados na literatura.

As Seções 4 e 5 abordam, respectivamente, os efeitos do aumento do número de classes e da ordem de predição no CCDC. Na Seção 6 são vistas as vantagens em se explorar a correlação temporal no estabelecimento da regra de predição do CCDC. A Seção 7 trata dos efeitos da tabela de códigos de Huffman sobre a taxa de compactação. Finalmente, a Seção 8 encerra o artigo, com uma síntese das principais conclusões do trabalho.

2. DESCRIÇÃO DOS CODIFICADORES DE CADEIA DIFERENCIAIS

2.1 O CODIFICADOR DE CADEIA DIFERENCIAL

O codificador de cadeia diferencial é o tipo básico do codificador de forma sem perda baseado em contornos. O seu princípio fundamental consiste em seguir o contorno de um objeto e codificar a direção da localização do próximo pixel de fronteira pela diferença em relação à direção anterior [2]. As várias versões do codificador consideram o pixel tendo 4 ou 8 vizinhos, em grades retangulares, ou 6 em grades hexagonais. Uma das mais conhecidas propostas de codificador de cadeia foi apresentada por Freeman [4], tendo provocado vários estudos posteriores [5]-[9]. A curva de contorno é considerada como contínua. No caso de grade retangular com 4 ou 8 vizinhos, a localização do próximo pixel, indicada pela direção diferencial, pode ser representada por 2 ou 3 bits, respectivamente, o que resulta em taxas médias de 1,2 ou 1,4 bits/pixel de fronteira para codificação sem perda [10]. A Figura 1 exemplifica o algoritmo.

Algumas das variações propostas com o objetivo de aumentar a eficiência de codificação incluem: (1) códigos de cadeia generalizados, com comprimentos de ligações e resoluções angulares adicionais [5]; (2) utilização de padrões na seqüência de códigos para aumento da eficiência de codificação [8]; (3) emprego da dependência estatística entre ligações sucessivas [9]; (4) simplificação do contorno para aumentar a eficiência de codificação [11], [12].

2.2 O CODIFICADOR DE CADEIA DIFERENCIAL CONDICIONAL (CCDC)

O codificador de cadeia diferencial proposto por Chung *et alii* [3] incorpora um aperfeiçoamento àquele descrito na Seção 2.1. Ele explora as probabilidades de ocorrência das direções diferenciais d_i , isto é, as possíveis mudanças de direção no i -ésimo pixel do contorno, condicionadas à direção diferencial do pixel anterior d_{i-1} .

Como ilustrado na Fig. 1, num codificador de cadeia diferencial as mudanças de direção entre os pixels do contorno podem assumir os valores $\{-3, -2, -1, 0, 1, 2, 3, 4\}$ quando a grade é retangular com 8 vizinhos. Em [3], essas direções diferenciais são reunidas em 2 classes ($A = \{0\}$ e $B = \{-3, -2, -1, 1, 2, 3, 4\}$) ou em 3 classes ($A = \{0\}$, $B = \{-3, -2, -1\}$ e $C = \{1, 2, 3, 4\}$). Estimam-se, então, a partir de um conjunto de imagens de teste, as probabilidades condicionais $P(d_i/A_{i-1})$ - probabilidade de

que o i -ésimo pixel do contorno apresente direção diferencial d_i , dado que a direção diferencial do $(i-1)$ -ésimo pixel pertence à classe A-, $P(d_i/B_{i-1})$ - probabilidade de que o i -ésimo pixel do contorno apresente direção diferencial d_i , dado que a direção diferencial do $(i-1)$ -ésimo pixel pertence à classe B -, e, se for o caso, $P(d_i/C_{i-1})$ - probabilidade de que o i -ésimo pixel do contorno apresente direção diferencial d_i , dado que a direção diferencial do $(i-1)$ -ésimo pixel pertence à classe C.

A partir da transmissão das condições iniciais (coordenadas do primeiro pixel, direção absoluta do primeiro para o segundo pixel e direção diferencial do segundo para o terceiro pixel), obtém-se, para cada pixel, a predição da correspondente direção diferencial a partir da direção diferencial do pixel anterior. A informação a ser codificada é o erro de predição da direção diferencial, isto é, a diferença entre as direções diferenciais real e predita, dado por

$$e_i = d_i - \hat{d}_i, \quad (1)$$

sendo a direção predita é obtida como

$$\hat{d}_i = \operatorname{argmax}_{d_i} P(d_i | \text{Classe de } d_{i-1}). \quad (2)$$

Um codificador por entropia (Huffman ou aritmético adaptativo) é, então, empregado para codificar o erro de predição.

Os resultados apresentados para o CCDC indicam uma melhora na eficiência de codificação em relação ao codificador de cadeia diferencial sem predição.

3. IDENTIFICAÇÃO DE POSSÍVEIS MUDANÇAS PARA O AUMENTO DE DESEMPENHO DE UM CODIFICADOR DE CADEIA

No extenso trabalho de avaliação de codificadores de forma realizado pelo grupo de elaboração do padrão MPEG-4 [13], concluiu-se que, para a codificação sem perdas, o codificador de cadeia diferencial (CCD), empregado pelo Codificador Baseado em Vértice [14], é o mais eficiente para o modo intraquadro. No modo interquadro, a codificação baseada em mapa de bits apresentou melhor desempenho. Neste modo, o Codificador Baseado em Vértice se vale da predição temporal, determinando e codificando o erro entre os contornos de dois quadros sucessivos. Entretanto, o seu resultado inferior no modo interquadro motivou, neste trabalho, a investigação de possíveis alterações do CCD que, a partir de uma maior exploração da correlação temporal, aumentassem o seu desempenho nesse modo. Por outro lado, o CCDC proposto em [3] elevou o desempenho do CCD com a exploração da correlação espacial entre os pixels de contorno vizinhos. Buscou-se, então, neste trabalho explorar ainda mais que em [14] a correlação temporal e mais que em [3] a correlação espacial, quando possível de uma forma integrada. Foram avaliados os efeitos de diversas alterações no codificador de cadeia diferencial condicional para, finalmente, concluir-se sobre o esquema mais eficiente. Ampliou-se o número de classes de predição espacial e foram empregados os dois pixels anteriores do contorno, como contexto de predição espacial, em vez de apenas um. Quanto à correlação temporal, empregaram-se

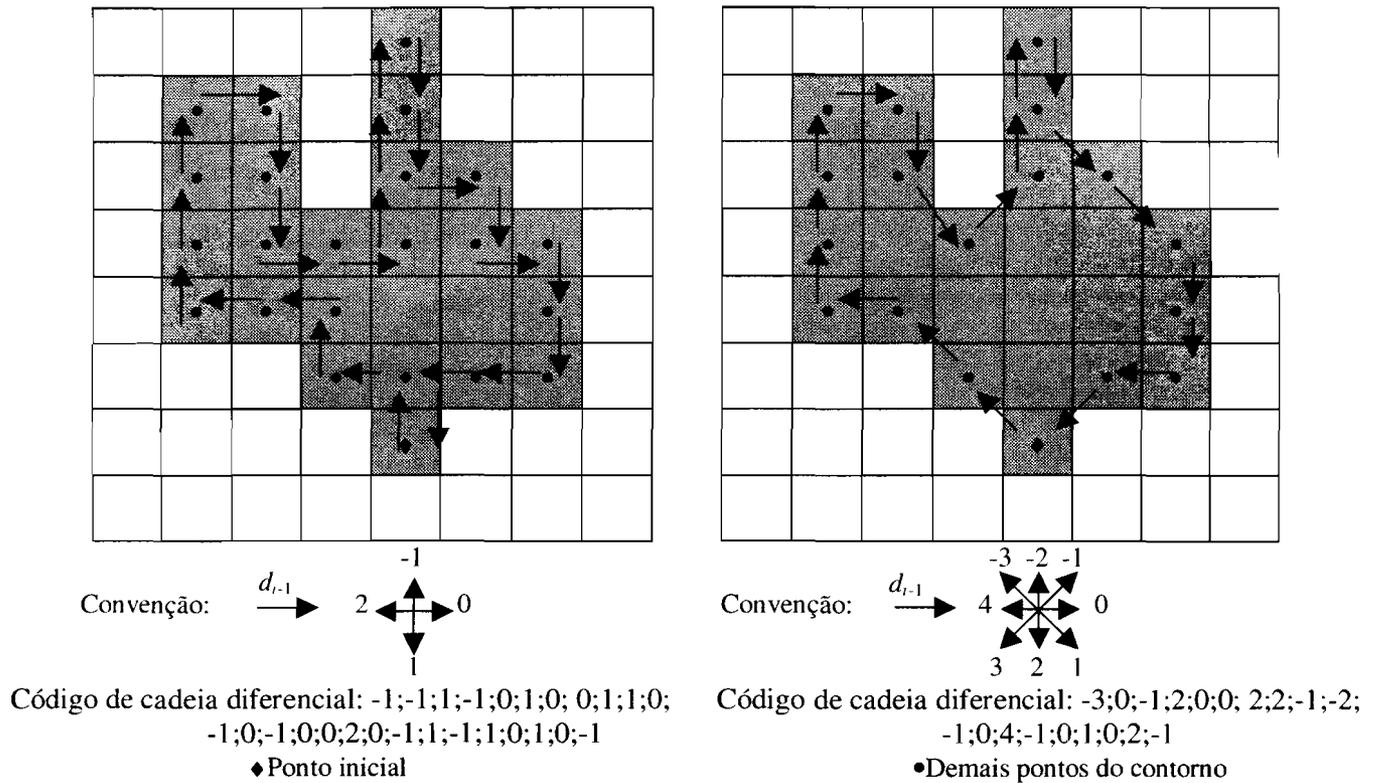


Figura 1. Exemplo de Código de Cadeia.

as frequências relativas do quadro anterior para estimar, adaptativamente, as distribuições de probabilidade condicional de cada quadro. A descrição detalhada das propostas e os resultados obtidos são apresentados nas seções seguintes.

4. EFEITO DO AUMENTO DO NÚMERO DE CLASSES NO CCDC

Em [3], mostrou-se que o emprego de três classes de predição em vez de apenas duas, resulta em melhor desempenho. De modo a se determinar o limite do aumento de desempenho em função do aumento do número de classes, foram realizados experimentos em que os contornos de objetos foram codificados empregando-se 3, como em [3], e 5 classes de predição. Os objetos utilizados nos experimentos foram os pertencentes às imagens Peppers (2 objetos) e Lena (1 objeto) apresentados no Apêndice A. No caso de 5 classes, foram empregadas as seguintes: $C_{1,i} = \{0\}$; $C_{2,i} = \{-1\}$; $C_{3,i} = \{1\}$; $C_{4,i} = \{-3, -2\}$; e $C_{5,i} = \{2, 3, 4\}$.

Inicialmente, utilizou-se ordem de predição igual a 1, como em [3], isto é, a predição de d_i baseou-se apenas na classe de d_{i-1} . As regras empregadas na predição, determinadas a partir das frequências relativas determinadas sobre os três objetos utilizados, são apresentadas a seguir.

- 3 classes:

- se a classe de d_{i-1} é A, atribuir 0 a d_i ;
- se a classe de d_{i-1} é B, atribuir +1 a d_i ;
- se a classe de d_{i-1} é C, atribuir -1 a d_i .

- 5 classes:

- se a classe de d_{i-1} é $C_{1,i}$, atribuir 0 a d_i ;
- se a classe de d_{i-1} é $C_{2,i}$, atribuir +1 a d_i ;
- se a classe de d_{i-1} é $C_{3,i}$, atribuir -1 a d_i ;
- se a classe de d_{i-1} é $C_{4,i}$, atribuir +2 a d_i ;
- se a classe de d_{i-1} é $C_{5,i}$, atribuir -1 a d_i .

Os erros de predição foram codificados por Huffman, com comprimento de símbolos igual a 1, empregando-se distribuição de probabilidade otimizada para cada tipo de imagem. Os resultados obtidos, em termos do número de bits/pixel resultante, constam da Tabela 1. Esses resultados não incluem os bits necessários à codificação das informações iniciais (coordenadas do primeiro pixel, direção absoluta do segundo e direção diferencial do terceiro) e da tabela de códigos de Huffman. Esse acréscimo de informação transmitido é idêntico para os dois métodos e sua omissão não tem qualquer reflexo sobre a comparação e a conclusão a respeito de qual é o mais eficiente.

Nº de Classes	Desempenho (bits/pixel)		
	Pepper (Obj1)	Pepper (Obj2)	Lena
3	1,3658	1,5271	1,5565
5	1,3687	1,5396	1,5425

Tabela 1. Resultados obtidos com ordem de predição 1.

Os resultados permitem concluir que o aumento do número de classes de 3 para 5 não é vantajoso em termos de eficiência

de codificação. Tal comportamento é explicado por dois fatores: a reduzida frequência de ocorrência das direções diferenciais $-3, -2, 2, 3$ e 4 ; e a semelhança das regras de predição das classes $B = \{-3, -2, -1\}$, $C_{2,i} = \{-1\}$ e $C_{4,i} = \{-3, -2\}$ e das classes $C = \{1, 2, 3, 4\}$, $C_{3,i} = \{1\}$ e $C_{5,i} = \{2, 3, 4\}$. Deve-se notar que $B = C_{2,i} \cup C_{4,i}$ e $C = C_{3,i} \cup C_{5,i}$.

De modo a se poder avaliar o efeito combinado dos aumentos do número de classes e da ordem de predição, foi realizada a codificação empregando predição com 3 e 5 classes, respectivamente, a partir das 9 e 25 possíveis combinações de classes dos dois pixels anteriores. Os resultados são apresentados na Tabela 2 e não incluem, pelas mesmas razões do caso anterior, os bits necessários à codificação das informações iniciais e da tabela de códigos de Huffman. Esses resultados confirmam que o aumento de três para cinco classes de predição não resulta em melhora de eficiência de codificação.

Nº de Classes	Desempenho (bits/pixel)		
	Pepper (Obj1)	Pepper (Obj2)	Lena
3	1,3400	1,5139	1,4764
5	1,3490	1,5180	1,4754

Tabela 2. Resultados obtidos com ordem de predição 2.

Tendo em vista esses resultados, em todos os experimentos seguintes foram consideradas apenas três classes de predição (as classes A, B e C do CCDC).

5. EFEITO DO AUMENTO DA ORDEM DE PREDIÇÃO NO CCDC

Nesta seção será avaliado o efeito isolado do aumento da ordem de predição no método CCDC, utilizando-se um ou os dois pixels anteriores para a predição de d_i , a direção diferencial do i -ésimo pixel. Os métodos examinados foram os seguintes:

1. CCD: sem qualquer predição, consistindo na codificação das direções diferenciais;
2. CCDC31: o CCDC com ordem de predição igual a 1, conforme em [3], com número de classes igual a 3. A regra de decisão foi derivada da distribuição de probabilidade global das imagens de teste, ao contrário do experimento anterior, em que as distribuições utilizadas eram específicas de cada objeto avaliado;
3. CCDC32: o CCDC com ordem de predição igual a 2. Também, como no caso do CCDC31, a regra de decisão foi derivada da distribuição de probabilidade global das imagens de teste;
4. CCDCA: o CCDC com ordem de predição igual a 2 e regra de decisão derivada adaptativamente da própria imagem processada. A distribuição de probabilidade e, em decorrência, a regra de decisão empregada, são atualizadas dinamicamente após o processamento de cada pixel e aplicadas ao seguinte. Esta é uma das estruturas que está sendo proposta neste trabalho.

A regra de decisão do método CCDC31 é dada pela equação (1). As dos métodos CCDC32 e CCDCA são dadas por

$$\hat{d}_i = \underset{d}{\operatorname{argmax}} P(d_i | \text{Classe de } d_{i-1}, \text{ Classe de } d_{i-2}). \quad (3)$$

O que varia entre as regras de decisão do CCDC32 e do CCDCA é a forma de determinação das probabilidades condicionais. Para o primeiro método, foram consideradas as frequências relativas globais referentes aos três objetos processados, para as nove possibilidades de combinação das classes de d_{i-1} e d_{i-2} . No método CCDCA, a frequência relativa é obtida adaptativamente das próprias direções diferenciais decodificadas. À medida em que a direção diferencial relativa a um determinado pixel de contorno é decodificada no receptor, ela é utilizada para atualizar as frequências relativas utilizadas para derivar a regra de decisão. Essas frequências relativas são referentes a todas as direções diferenciais de pixels de contorno do objeto sendo processado já decodificadas.

As direções diferenciais do método CCD e os erros de predição dos demais métodos foram codificados por Huffman, com comprimentos de símbolos conjuntamente codificados iguais a 1, 2 e 3. A tabela das palavras-código foi otimizada para cada objeto.

Os resultados obtidos são apresentados nas Tabelas 3 e 4. Não incluem os bits necessários à codificação das coordenadas do primeiro pixel, da direção absoluta do primeiro para o segundo pixel e da tabela de códigos de Huffman, visto que o acréscimo de informação transmitido é idêntico para todos os métodos. Não incluem também os bits necessários para transmitir as regras de decisão. Isso pode implicar um acréscimo adicional para os métodos CCDC31 e CCDC32. Esse efeito será examinado posteriormente.

Contata-se claramente desses resultados que o aumento da ordem de predição de 1 (CCDC31) para 2 (CCDC32) eleva o desempenho do codificador. O método CCDCA, um dos propostos neste artigo, apresenta desempenho levemente inferior ao do CCDC32, mas pode se constituir numa alternativa interessante caso tenha que se transmitir a regra de decisão no CCDC32, o que aumentaria a respectiva taxa. Como no CCDCA essa regra é derivada da própria imagem, não haveria esse acréscimo.

Em todos os casos, o aumento do número de símbolos codificados conjuntamente por Huffman elevou a eficiência de codificação quando não se considera o possível acréscimo de transmissão da tabela de códigos. O efeito da transmissão da tabela de códigos também será examinado posteriormente. Cabe ressaltar que para os casos de codificação conjunta de 2 e 3 símbolos, o método CCDC31 mostrou-se inferior ao CCD, ao contrário do apresentado em [3].

6. USO DA CORRELAÇÃO TEMPORAL NA REGRA DE PREDIÇÃO DO CCDC

Esta seção avalia a possível vantagem em se considerar a correlação temporal entre quadros de uma seqüência de imagens no estabelecimento da regra de decisão empregada na predição do CCDC. Para que fossem efetivas as comparações por meio das taxas em bits/pixel, foram considerados, nos

Método	Nº de Classes	Ord. de Pred.	Desempenho (bits/pixel)								
			Pepper (Obj1)			Pepper (Obj2)			Lena		
			Nº de símb/Huffman			Nº de símb/Huffman			Nº de símb/Huffman		
			1	2	3	1	2	3	1	2	3
CCD	-	-	1,6303	1,3106	1,1407	1,6992	1,4024	1,3200	1,6214	1,3621	1,2420
CCDC31	3	1	1,5661	1,3877	1,2360	1,5976	1,4546	1,3278	1,5623	1,3923	1,3042
CCDC32	-	2	1,3930	1,1557	1,1050	1,5627	1,3381	1,3068	1,4863	1,2637	1,2176
CCDCA	-	2	1,4068	1,1769	1,1291	1,6149	1,4834	1,3854	1,4875	1,2902	1,2420

Tabela 3. Resultados por imagem para o aumento da ordem de predição.

Método	Nº de Classes	Ord. Pred.	Desempenho (bits/pixel)		
			Nº de símb/Huffman		
			1	2	3
CCD	-	-	1,6434	1,3586	1,2349
CCDC31	3	1	1,5722	1,4068	1,2921
CCDC32		2	1,4809	1,2539	1,2103
CCDCA		2	1,4983	1,3090	1,2484

Tabela 4. Resultados médios obtidos com o aumento da ordem de predição.

seus respectivos cálculos, os bits necessários para a transmissão das regras de decisão otimizadas para cada quadro, no caso do método denominado Código de Cadeia Diferencial Condicionada ao Quadro Corrente (CCDCQC), bem como das condições iniciais (coordenadas do primeiro pixel, direção absoluta do primeiro para o segundo pixel e uma ou as duas direções diferenciais seguintes, quando aplicável) para todos os métodos. A Tabela 5 apresenta os dados relativos ao número de bits necessário para codificar as condições iniciais e a regra de decisão em cada método empregado na análise. Contém, ainda, o número de direções diferenciais e/ou o número de erros de predição transmitidos em cada método. É importante ressaltar que os métodos que usam a predição da direção diferencial empregam três classes.

6.1 MÉTODOS EMPREGADOS NA ANÁLISE

O primeiro método é o CCD. Nota-se da Tabela 5 que o acréscimo de informação transmitida é de apenas 19 bits, correspondentes às coordenadas do primeiro pixel e à direção absoluta do segundo. Essas informações são necessárias para identificar o ponto inicial do contorno e o referencial para a primeira direção diferencial transmitida, entre o segundo e o terceiro pixels.

Para os métodos CDCC31 [3] e CDCC32 as regras de decisão foram inferidas de duas seqüências de treinamento (Children e Coast Guard) e a avaliação foi procedida com outras duas seqüências de teste (Fish-and-Logo e Weather), conforme será detalhado posteriormente. Quadros das quatro seqüências utilizadas são mostrados no Apêndice A. Além das mesmas informações iniciais necessárias para o método CCD, no caso do CCDC31 há, ainda, a necessidade de se informar a direção diferencial do segundo para o terceiro pixel. Essa informação serve de base para a primeira predição de direção diferencial, referente à direção entre o terceiro e o quarto pixel. Para o CCDC32, deve-se transmitir inicial-

mente duas direções diferenciais, visto que este método emprega ordem de predição igual a 2. O método do Código de Cadeia Diferencial Condicionada ao Quadro Corrente (CCDCQC) é uma modalidade de CCDC32, com as regras de decisão otimizadas para cada quadro. Nesse caso, as regras de decisão têm de ser transmitidas, acarretando um aumento adicional médio da informação transmitida, em relação ao CCDC32, de 12 bits. A alocação desses 12 bits pode ser compreendida com o exemplo da Tabela 6. As duas primeiras colunas referem-se às classes das duas direções diferenciais anteriores empregadas na predição. A terceira coluna exemplifica as direções preditas para cada caso, a partir dos valores máximos das probabilidades condicionais. A última coluna mostra a alocação de códigos às direções preditas. Cabe ressaltar que, em todos os casos levantados, das nove direções preditas, pelo menos 5 sempre foram iguais a 0 e as demais iguais a +1 ou -1. Isso permitiu, com um código de um bit alocado a 0 e com códigos de 2 bits alocados a +1 e a -1, que, no máximo, 13 bits e, na média, 12 bits fossem necessários para transmitir a regra de predição. Esse tipo de estrutura, empregando e transmitindo a regra de predição otimizada e codificada, foi proposto neste trabalho.

O último método, o do Código de Cadeia Diferencial Condicionada ao Quadro Anterior (CCDCQA), é outra estrutura proposta neste trabalho. Consiste também numa versão de CCDC32, porém as regras de decisão são derivadas da freqüências relativas do quadro anterior. Nesse caso, não há necessidade de transmiti-las, pois o receptor já disporá do quadro anterior ao receber o corrente. Note-se que este método explora a correlação temporal entre quadros sucessivos.

Em todos os métodos, foram exploradas, como anteriormente, três possibilidades de codificação por Huffman, com a tabela de códigos otimizada para cada seqüência de 1 símbolo isolado, bem como de 2 e de 3 símbolos conjuntamente. A redução da taxa de transmissão devido à possível transmissão da tabela de códigos não foi levada em conta. Esse efeito será considerado posteriormente. Uma vez mais, é importante ressaltar que isso não interfere com a comparação de desempenho entre os métodos, visto que o acréscimo da informação transmitida é o mesmo para todos.

6.2 RESULTADOS

Foram empregadas para treinamento as imagens dos quadros 1 a 11 da seqüência Children, sendo considerados dois objetos (os dois meninos) em cada quadro, totalizando 5737 pixels, e dos quadros 1 a 14 da seqüência Coast Guard,

Método	Cabeçalho dos objetos transmitidos em cada quadro, contendo as condições iniciais e a regra de decisão							Cadeia Diferencial codificada por Huffman (pixels)	
	Sem codificação (bits)					Codificado por Huffman (pixels)			
	Coordenada horizontal do primeiro pixel x_0	Coordenada vertical do primeiro pixel y_0	Direção absoluta do segundo pixel r_0	Regra de decisão	Acréscimo total de informação transmitida em bits	Direção diferencial do terceiro pixel	Direção diferencial do quarto pixel	Nº de direções diferenciais transmitidas	Nº de erros de direções diferenciais transmitidos
CCD	8	8	3	-	19	-	-	N-2	-
CCDC31	8	8	3	-	19	1	-	-	N-3
CCDC32	8	8	3	-	19	1	1	-	N-4
CDCCQC	8	8	3	12	31	1	1	-	N-4
CDCCQA	8	8	3	-	19	1	1	-	N-4

Obs.: N é o número de pixels do contorno do objeto.

Tabela 5. Dados numéricos da codificação.

Classe de d_{i-2}	Classe de d_{i-1}	Direção predita	Código transmitido
B (-)	B (-)	+1	10
	A (0)	0	0
	C (+)	0	0
A (0)	B (-)	+1	10
	A (0)	0	0
	C (+)	0	0
C (+)	B (-)	0	0
	A (0)	0	0
	C (+)	-1	11

Tabela 6. Exemplo de alocação de bits à regra de predição.

igualmente com dois objetos (as duas embarcações) em cada quadro, totalizando 2700 pixels. Da estatística dos contornos desses objetos foram derivadas as regras de predição dos métodos CCDC31 e CCDC32.

Foram empregadas para teste as imagens dos quadros 9 a 23 da seqüência Fish-and-Logo, sendo considerados três objetos (os 3 peixes) em cada quadro, totalizando 13393 pixels, e dos quadros 1 a 38 da seqüência Weather, com um objeto (a locutora) em cada quadro, totalizando 13196 pixels. A Tabela 7 resume os resultados médios obtidos para as seqüências de teste.

Pode-se constatar que, dentre os métodos investigados, o CCDCQA é o que apresenta o melhor desempenho, o que permite concluir que a exploração da correlação temporal entre quadros possibilita reduções na taxa de transmissão do código de cadeia diferencial. Há, porém, que se ter um cuidado especial quando o CCDCQA for empregado: o método é sensível a erros introduzidos pelo canal. Dessa forma, deve ser estabelecida uma "taxa de reinicialização", isto é, a cada determinado número de quadros transmitidos deve-se adotar, simultaneamente no transmissor e no receptor, novamente a regra de decisão de inicialização do sistema.

Analisando-se os resultados isoladamente para cada tipo de seqüência, é possível constatar, também, que o ganho do método CCDCQA em relação ao CCD é maior nos con-

tornos mais simples (aqueles em que preponderam as pequenas mudanças de direção - $d_i = \pm 1$ - ou nenhuma - $d_i = 0$) e nos que variam mais lentamente no tempo, já que nesses casos a correlação espacial e a temporal são mais intensas. É o caso da seqüência Weather, para a qual o ganho em desempenho do CCDCQA foi mais acentuado que para a seqüência Fish-and-Logo, cujos contornos são mais complexos e variáveis no tempo.

Método	Nº de Classes	Ordem de Predição	Desempenho médio (bpp)		
			Nº de símb/Huffman		
			1	2	3
CCD	-	-	1,62	1,34	1,19
CCDC31	3	1	1,59	1,36	1,20
CCDC32		2	1,56	1,31	1,18
CCDCQC		2	1,57	1,33	1,19
CCDCQA		2	1,55	1,30	1,17

Tabela 7. Resultados de desempenho.

7. EFEITO DA TABELA DE CÓDIGOS DE HUFFMAN SOBRE A TAXA DE COMPRESSÃO

Esta seção avalia os efeitos sobre a taxa de compressão de cada uma das possíveis escolhas da tabela de códigos empregada para a codificação por Huffman. O uso da tabela otimizada para o objeto traz ganhos evidentes em compressão na codificação. Entretanto, exige a transmissão da tabela, o que reduz aqueles ganhos. Outra possibilidade é a utilização de uma tabela derivada de um conjunto de imagens de treinamento. Os efeitos, neste caso, são opostos ao anterior? a eficiência de compressão obtida com o uso da tabela é menor, mas não é necessária a transmissão da tabela de códigos. Existe ainda uma terceira alternativa que consiste no emprego da tabela de códigos do quadro anterior, explorando assim a

possível correlação temporal entre essa tabela e a otimizada para a quadro atual. Como o receptor já dispõe a priori das direções diferenciais do quadro anterior, pode derivar a tabela de códigos das frequências relativas dessas direções, sem a necessidade de qualquer acréscimo da informação a ser transmitida. Assim, foram avaliados os efeitos sobre a taxa de compressão das três seguintes possibilidades:

- tabela de códigos otimizada para o objeto considerado, com transmissão da tabela e conseqüente redução da taxa de transmissão;
- tabela de códigos obtida de amostras de treinamento (*ensemble*);
- tabela de códigos correspondente ao quadro anterior.

7.1 MÉTODOS EMPREGADOS NA ANÁLISE

Como a avaliação da Seção 6 indicou o método CCDCQA como o mais eficiente, os estudos nesta seção foram restritos a ele e ao CCD, tomado como parâmetro de comparação. Para a segunda possibilidade - tabela de códigos obtida de amostras de treinamento - as tabelas foram derivadas das seqüências Children (quadros 1 a 11) e Coast Guard (quadros 1 a 14). Para as três possibilidades de tabela, a avaliação foi procedida com as seqüências de teste - Fish-and-Logo (quadros 9 a 23) e Weather (quadros 1 a 38).

Quanto ao número de símbolos conjuntamente codificados por Huffman, foram avaliadas, a exemplo de análises anteriores, três alternativas: 1, 2 ou 3 símbolos.

No tocante ao emprego das tabelas de códigos otimizadas para cada quadro, cabe observar que a transmissão não codificada dessas tabelas representaria um acréscimo da informação transmitida muito elevado, o que inviabilizaria a possibilidade de obter ganho em compressão com as suas transmissões. Para o caso de 1 símbolo por código seriam necessários, em média, 35 bits ($35/8 = 4,375$ bits/código), conforme exemplifica a Tabela 8.

Para os casos de 2 e 3 símbolos por código, as quantidades médias de bits necessários seriam da ordem de 267 ($267/64 = 4,172$ bits/código) e 2036 ($2036/512 = 3,977$ bits/código), respectivamente. Tendo em vista esses números, estudou-se uma maneira de transmitir a tabela codificada com um menor número de bits, através de um procedimento de quantização das tabelas. Do exame de todas as possíveis tabelas obtidas com as imagens de treinamento e de teste, foram identificadas apenas 5 variações para a tabela correspondente a 1 símbolo por código. Para o caso de 2 símbolos por código, somente 52 variações foram identificadas. Portanto, 3 bits foram suficientes para codificar as possíveis tabelas para comprimento 1 e 6 bits para as tabelas correspondentes a comprimento 2. Para analisar o caso de 3 símbolos conjuntamente codificados (são 512 entradas na tabela, que podem combinar-se de maneira bastante variada), adotou-se, por similaridade com os outros dois casos, 9 bits para codificar a tabela.

No processo de quantização das tabelas, sempre que não há uma tabela codificada (TC) correspondente à otimizada (TO) para o quadro (aqui isso ocorreu apenas para o caso de 3 símbolos conjuntamente codificados), escolhe-se a tabela mais próxima. O critério utilizado para essa escolha é o

Direção Diferencial	Código
-3	1111110
-2	1110
-1	110
0	0
1	10
2	11110
3	111110
4	1111111
Total de bits	35

Tabela 8. Exemplo de tabela de Huffman para 1 símbolo por código.

seguinte: para cada seqüência de símbolos (ou direções diferenciais) a serem conjuntamente codificados, subtrai-se o número de bits do código em TO do correspondente número de bits do código em TC e multiplica-se o resultado pela frequência de ocorrência da seqüência considerada; a soma de todos esses valores parciais corresponde à distância entre as tabelas, em termos de comprimento médio de código, devendo ser selecionada a tabela codificada que apresentar menor distância da otimizada.

A Tabela 9 apresenta a distribuição de bits auxiliares para os dois métodos empregados nesta análise, apresentando o acréscimo devido à transmissão das tabelas de códigos de Huffman para os casos de 1, 2 ou 3 símbolos conjuntamente codificados.

7.2 RESULTADOS

A Tabela 10 apresenta os resultados médios correspondentes aos métodos CCD e CCDCQA obtidos para as seqüências de teste.

Os melhores desempenhos foram obtidos para o método CCDCQA com transmissão codificada das tabelas de Huffman. Cabe observar que o emprego de uma tabela não otimizada para o quadro quando são codificados 2 e, sobretudo, 3 símbolos conjuntamente, degrada consideravelmente o desempenho. Isso se explica pelo fato de que muitas palavras-código, nesse caso, são longas. A atribuição de uma palavra-código muito mais longa que a ótima a uma seqüência de símbolos relativamente freqüente pode provocar um aumento significativo da taxa medida em bits/pixel. Ao contrário do que ocorreu com a regra de predição, a exploração da correlação temporal pelo emprego da tabela de códigos de Huffman do quadro anterior não surtiu efeitos benéficos sobre o desempenho de codificação.

8. CONCLUSÕES

Neste artigo foram consideradas diversas modificações no codificador de cadeia diferencial condicional para compactação da informação de forma em codificação de vídeo baseada em objeto. Elas exploram de modo mais acentuado as correlações espacial e temporal. Além disso, ao contrário do que foi feito em [3], foram considerados os efeitos sobre a taxa de transmissão das diversas informações

Método	Cabeçalho dos objetos transmitidos em cada quadro, contendo as condições iniciais e a regra de decisão						Cadeia Diferencial codificada por Huffman (pixels)		
	Sem codificação (bits)					Codificado por Huffman (pixels)		Nº de direções diferenciais transmitidas	Nº de erros de direções diferenciais transmitidos
	Coordenada horizontal do primeiro pixel x_0	Coordenada vertical do primeiro pixel y_0	Direção absoluta do segundo pixel r_0	Tabela de Códigos de Huffman	Acréscimo total em bits	Direção diferencial do terceiro pixel	Direção diferencial do quarto pixel		
CCD - Tab. Otimizada	8	8	3	3/6/9	22/25/28	-	-	N-2	-
CCD - Tab. do Ensemble	8	8	3	-	19	-	-	N-2	-
CCD - Tab. Quadro Anterior	8	8	3	-	19	-	-	N-2	-
CCDCQA - Tab. Otimizada	8	8	3	3/6/9	22/25/28	1	1	-	N-4
CCDCQA - Tab. do Ensemble	8	8	3	-	19	1	1	-	N-4
CCDCQA - Tab. Quadro Anterior	8	8	3	-	19	1	1	-	N-4

Tabela 9. Dados numéricos da codificação.

Método	Variação do Método	Desempenho médio para os objetos de teste (bits/pixel)		
		Nº de símbolos/ Huffman		
		1	2	3
CCD	Indicação Tabela Huffman.	1,63	1,36	1,21
CCD	Tab. Huffman. ensemble	1,64	1,40	1,82
CCD	Tab. Huffman quadro ant.	1,63	1,38	1,71
CCDCQA	Indicação Tab. Huffman	1,56	1,32	1,19
CCDCQA	Tab. Huffman ensemble	1,59	1,44	2,29
CCDCQA	Tab. Huffman quadro ant.	1,58	1,41	2,11

Tabela 10. Desempenho Médio Obtido pela Variação da Tabela de Huffman.

preliminares que devem ser enviadas ao receptor. Tais informações incluem sempre as coordenadas do primeiro pixel do contorno e a direção absoluta do segundo pixel. Dependendo do método, pode ser necessário ainda transmitir a direção diferencial do segundo para o terceiro e do terceiro para o quarto pixels, a regra de predição e a tabela de códigos de Huffman. O fato de não se considerar todos esses acréscimos de informação transmitida pode alterar significativamente os resultados relativos e as conclusões sobre a eficiência relativa dos métodos.

Constatou-se que o aumento de três para cinco no número de classes empregadas na predição da direção diferencial não traz ganhos em termos de eficiência de codificação, ao contrário do que ocorre quando se aumenta de duas para três classes.

A utilização dos dois pixels de contorno anteriores, ao invés de apenas um, proporciona uma efetiva melhora nos resultados da predição espacial e, em consequência, no aumento da eficiência de codificação. Da mesma forma, a

exploração da correlação temporal entre quadros sucessivos, pela derivação da regra de decisão empregada na predição a partir do quadro anterior, melhora o desempenho do codificador. Adicionalmente, o desempenho melhora quando se eleva o número de símbolos conjuntamente codificados por Huffman de 1 para 2 e de 2 para 3, desde que se utilize a tabela de códigos associada ao objeto, mesmo tendo-se que levar em conta a transmissão de informações preliminares relativas a essa tabela. É importante salientar, porém, que esse resultado considera que a tabela não é necessariamente ótima, uma vez que emprega um procedimento de quantização para redução do número de bits adicionais transmitidos, conforme descrito na Seção 7 deste artigo.

Dos três métodos propostos - o CCDCA na Seção 5 e os CCDCQC e CCDCQA na Seção 6 -, o CCDCQA foi o que apresentou o melhor desempenho médio, superando os métodos CCD e CCDC31 encontrados na literatura.

Finalmente, concluiu-se que o ganho em desempenho obtido pelo emprego de predição da direção diferencial, explorando-se a correlação tanto espacial como temporal, varia com as características do contorno. Quanto mais simples o contorno (pequenas variações de direção) e quanto mais lenta for sua variação no tempo, melhor o desempenho, já que nesses casos a correlação espacial e a temporal entre os pixels do contorno são mais intensas. O método mais eficaz, o CCDCQA, o é sobretudo para imagens de contornos simples. Dessa forma, é uma alternativa bastante válida em aplicações em que a variação dos contornos dos objetos é pequena no tempo e no espaço, como em videoconferência, telefonia móvel, telejornais ou quando a taxa de amostragem espacial e a taxa de repetição de quadros forem elevadas.

APÊNDICE

A. IMAGENS EMPREGADAS NOS EXPERIMENTOS



Figura 2. Lena.

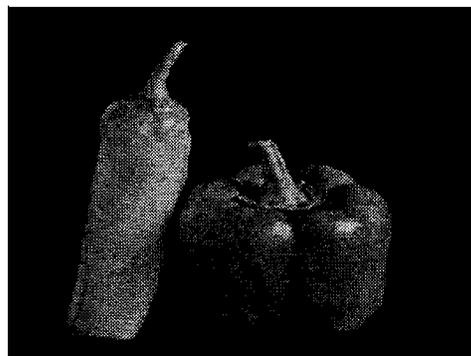


Figura 3. Peppers.

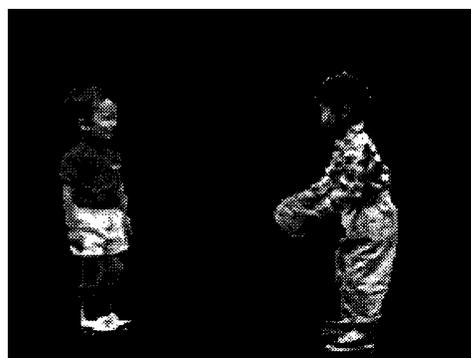


Figura 4. Children.

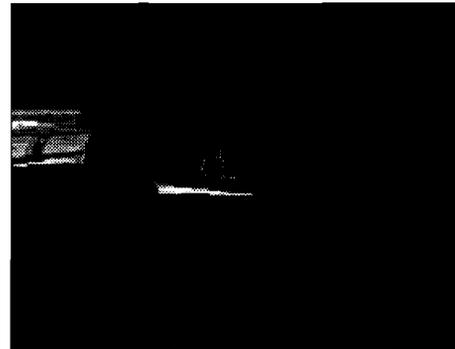
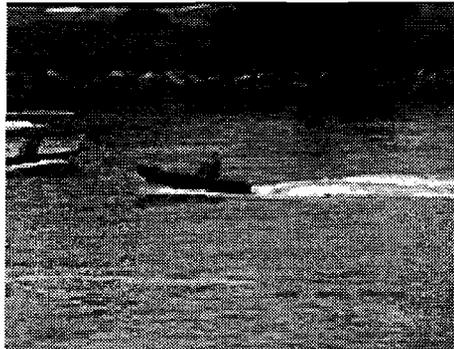


Figura 5. Cost Guard.

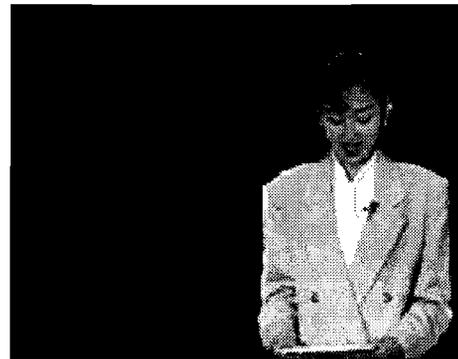


Figura 6. Weather.



Figura 7. Fish-and-Logo

REFERÊNCIAS

- [1] T. Sikora, "The MPEG-4 Video Standard Verification Model", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, n. 1, pp. 19-31, Feb. 1997.
 - [2] P. Salembier, F. Marques, and A. Gasull, "Coding of partition sequences", in *Video Coding*, L. Torres e M. Kunt, Eds. Boston, MA: Kluwer, 1996, pp. 125-170.
 - [3] J. Chun, J. Moon, and J. Kim, "Conditional differential chain coding for lossless representation of object contour", *Electronics Letters*, vol. 34, n. 1, pp. 55-56, Jan. 1998.
 - [4] H. Freeman, "On the encoding of arbitrary geometric configurations", *IRE Trans. Electron. Comput.*, vol. EC-10, pp. 260-268, June 1961.
 - [5] J. Sagri and H. Freeman, "Analysis of the precision of generalized chain codes for the representation of planar curves", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 533-539, Sept. 1981.
 - [6] J. Koplowitz, "On the performance of chain codes for quantization of line drawings", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 180-185, Mar. 1981.
 - [7] D. Neuhoff and K. Castor, "A rate and distortion analysis of chain codes for line drawings", *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 53-68, Jan. 1985.
 - [8] T. Kaneko and M. Okudaira, "Encoding of arbitrary curves based on the chain code representation", *IEEE Trans. Commun.*, vol. COM-33, pp. 697-707, July 1985.
 - [9] R. Prasad, J. W. Vieveen, J. H. Bons, and J. C. Arnbak, "Relative vector probabilities in differential chain coded line-drawings", in *Proc. IEEE Pacific Rim Conf. Communication, Computers and Signal Processing*, Victoria, Canada, June 1989, pp. 138-142.
 - [10] M. Eden and M. Kochen, "On the performance of a contour coding algorithm in the context of image coding. Part I: Contour segment coding", *Signal Process.*, vol. 8, pp. 381-386, July 1985.
 - [11] P. Nunes, P. Pereira, and F. Marques, "Multi-grid chain coding of binary shapes", in *Proc. Special Session on Shape Coding, ICIP97*, Santa Barbara, CA, 1997, pp. 114-117.
 - [12] T. Oselik and A. K. Katsaggelos, "Very low bit rate video coding based on statistical spatio-temporal prediction of motion, segmentation and intensity fields", in *Video Data Compression for Multimedia Computing*, H. H. Lin, S. Sun e H. Derin, Eds. Norwell, MA: Kluwer, 1997, pp. 313-353.
 - [13] A. K. Katsaggelos, L. P. Kondi, F. W. Meier, and G. M. Schuster, "MPEG-4 and rate-distortion-based shape-coding techniques", *Proc. IEEE*, vol. 86, n. 06, pp. 1126-1154, June 1998.
 - [14] K. J. O'Connell, "Object-adaptive vertex-based shape coding method", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 251-255, Feb. 1997.
- Emilio C. Acocella** graduou-se em engenharia eletrônica em 1989 e concluiu o mestrado em processamento de sinais em 1993, ambos realizados no Instituto Militar de Engenharia (IME). Concluiu o doutorado em processamento de sinais em 2000 no CETUC/PUC-Rio. Trabalha na área de processamento digital de imagem, área em que tem publicados artigos em congressos e revistas nacionais e internacionais. Foi professor do IME, da Escola Politécnica do Exército Equatoriano e do CEUB-Brasília. Dirige, atualmente, o Parque Regional de Manutenção/1, do Comando do Exército/Ministério da Defesa.
- Abraham Alcaim** recebeu o diploma de Engenheiro Eletricista e o título de Mestre em Ciências em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro (PUC/Rio) em 1975 e 1977, respectivamente, e os títulos de D.I.C. e Ph.D. pelo Imperial College of Science and Technology, University of London, em 1981. Desde 1976 ele é professor do Centro de Estudos em Telecomunicações da Universidade Católica (CETUC), tendo atualmente o cargo de Professor Associado. O Dr. Alcaim trabalha há mais de 25 anos nas áreas de processamento digital de voz e imagem. Ele é autor de diversos artigos publicados em congressos e revistas nacionais e internacionais. Em 1984 ele esteve por um período curto com o Centre National d'Etudes des Télécommunications (CNET), em Lannion, França, onde trabalhou em medidas de qualidade objetivas e subjetivas para codificadores de voz. De dezembro de 1991 a setembro de 1993 ele foi Cientista Visitante no Centro Científico Rio da IBM Brasil, onde trabalhou no projeto de novos codificadores de imagem, com aplicação especial para imagens obtidas por satélites de sensoriamento remoto. O Dr. Alcaim foi o Technical Program Chairman dos simpósios internacionais SBT/IEEE International Telecommunications Symposium de 1990 e 1994, e o Executive Chairman da IEEE Global Telecommunications Conference de 1999. Ele foi membro do Conselho Deliberativo da SBrT no período de 1996 a 2001 e membro do Comitê de Assessoramento de Engenharia Elétrica, Biomédica e Microeletrônica (CA-EE) do CNPq no período de 1998 a 2001. O Dr. Alcaim é o Editor da área de Processamento de Sinais da Revista da SBrT.