

Sequências de Baixa Discrepância Aplicadas à Inicialização do Algoritmo Linde-Buzo-Gray

E.A. LIMA¹, Escola Politécnica de Pernambuco, Universidade de Pernambuco, 50100-010, Recife, PE, Brasil.

P.H. ESPÍRITO SANTO², Universidade Federal de Pernambuco, 50740-530, Recife, PE, Brasil.

F. MADEIRO³, Escola Politécnica de Pernambuco, Universidade de Pernambuco, 50100-010, Recife, PE, Brasil.

Resumo. Sistemas de compressão de voz e imagem, esteganografia e reconhecimento de padrões estão entre possíveis aplicações da quantização vetorial (QV), cujo desempenho depende fortemente da qualidade dos dicionários projetados. Uma das técnicas mais amplamente utilizadas para o projeto de dicionários é o algoritmo LBG (Linde-Buzo-Gray). Neste artigo, as sequências de baixa discrepância são apresentadas como uma alternativa de inicialização do algoritmo. O desempenho da técnica apresentada é avaliado no cenário de compressão de imagem baseada em QV.

Palavras-chave. Quantização vetorial, algoritmo LBG, sequências de baixa discrepância.

1. Introdução

A Quantização Vetorial (QV) [5, 6] tem sido utilizada com sucesso em sistemas de codificação de voz [15, 18] e imagem [10, 13], permitindo elevadas taxas de compressão. Também tem sido usada em outras aplicações, tais como, identificação vocal [14], esteganografia e marca d'água digital [8, 1]. A QV pode ser vista como uma extensão da quantização escalar em um espaço multidimensional, encontra-se fundamentada na Teoria da Distorção Versus Taxa, formulada por Shannon, e é definida como um mapeamento Q de um vetor de entrada $\mathbf{x} \in \mathbb{R}^k$, em um subconjunto finito $W \subset \mathbb{R}^k$ contendo N vetores de reprodução. Assim, $Q : \mathbb{R}^k \rightarrow W$.

O dicionário $W = \{\mathbf{w}_i\}_{i=1}^N$ é o conjunto de vetores-código (vetores de reconstrução), k é a dimensão dos vetores-código e N é o tamanho do dicionário. O mapeamento Q particiona o espaço \mathbb{R}^k em N regiões disjuntas (também chamadas de células de Voronoi) S_i , $i = 1, 2, \dots, N$, em que $S_i = \{\mathbf{x} : Q(\mathbf{x}) = \mathbf{w}_i\}$, ou,

¹eal@poli.br.

²paulohugos@gmail.com

³madeiro@poli.br

alternativamente, $S_i = \{\mathbf{x} : d(\mathbf{x}, \mathbf{w}_i) \leq d(\mathbf{x}, \mathbf{w}_j), \forall j \neq i\}$, em que $d(\cdot, \cdot)$ é uma medida de distorção. O vetor-código \mathbf{w}_i é o vetor representante de todos os vetores de entrada pertencentes à célula S_i correspondente.

O desempenho dos sistemas de processamento de sinais baseados em QV depende fortemente dos dicionários projetados. Em sistemas de reconhecimento de locutor [14] baseado em QV, as taxas de identificação dependem dos dicionários projetados. Em sistemas de compressão de sinais baseados em QV, a qualidade de sinais reconstruídos depende dos dicionários projetados. O algoritmo Linde-Buzo-Gray (LBG) [12] constitui-se na técnica mais utilizada para o projeto de dicionários. Pode ser visto como o método das k -médias para clusterização de dados. A velocidade de convergência do algoritmo e o desempenho do dicionário final dependem do dicionário inicial [5, 12], razão pela qual estudos têm sido realizados com o objetivo de propor técnicas de inicialização do algoritmo LBG.

No presente artigo propõe-se o uso de sequências de baixa discrepância como alternativas para inicialização do algoritmo LBG. É realizada uma avaliação comparativa de desempenho com o algoritmo de inicialização apresentado por Katsavounidis et al. [9]. Os resultados de simulação também são comparados com os obtidos com uso de *splitting* [12] e *pairwise nearest neighbor* (PNN) [3, 4].

2. Algoritmos

2.1. Algoritmo LBG

Seja a iteração do algoritmo LBG denotada por n . Dados k , N e um limiar de distorção $\epsilon \geq 0$, o algoritmo LBG [12], também conhecido como GLA (*generalized Lloyd algorithm*), consiste da seguinte sequência de passos:

- *Passo 1)* inicialização: dado um dicionário inicial W_0 e um conjunto de treino $\mathbf{X} = \{\mathbf{x}_m; m = 1, 2, \dots, M\}$, faça $n = 0$ e $D_{-1} = \infty$;
- *Passo 2)* particionamento: dado W_n (dicionário na n -ésima iteração), aloque cada vetor de treino (vetor de entrada) na respectiva classe (célula de Voronoi) segundo o critério do vetor-código mais próximo; calcule a distorção

$$D_n = \sum_{i=1}^N \sum_{\mathbf{x}_m \in S_i} d(\mathbf{x}_m, \mathbf{w}_i) = \sum_{m=1}^M d(\mathbf{x}_m, Q(\mathbf{x}_m)), \quad (2.1)$$

em que $Q(\mathbf{x}_m) = \mathbf{w}_i$;

- *Passo 3)* teste de convergência (critério de parada): se $(D_{n-1} - D_n)/D_n \leq \epsilon$ pare, com W_n representando o dicionário final (dicionário projetado); caso contrário, continue;
- *Passo 4)* atualização do dicionário: compute os novos vetores-código como os centróides das classes de vetores; faça $W_{n+1} = W_n$; faça $n = n + 1$ e vá para *Passo 2*.

Em essência, no algoritmo LBG a função distorção decresce monotonicamente, uma vez que o dicionário é iterativamente atualizado visando satisfazer as condições de centróide e de vizinho mais próximo. Na dinâmica do algoritmo LBG, a distorção introduzida ao se representarem os vetores do conjunto de treinamento pelos correspondentes vetores-código (centróides) é monitorada a cada iteração. A regra de parada (teste de convergência) do algoritmo baseia-se nessa distorção monitorada – o treinamento do dicionário é encerrado quando $(D_{n-1} - D_n)/D_n \leq \epsilon$.

2.2. Algoritmo de distância máxima

Uma técnica de inicialização do algoritmo LBG foi proposta por Katsavounidis et al. [9]. Em essência, o método consiste em selecionar como vetor-código aquele que apresenta a **distância máxima** em relação aos outros vetores do dicionário. Seja um conjunto de treino $X = \{\mathbf{x}_m; m = 1, 2, \dots, M\}$. O algoritmo consiste nos seguintes passos:

Passo 1. Calcule as normas de todos os vetores do conjunto de treino X . O primeiro vetor-código (\mathbf{w}_1^0) de W_0 (dicionário inicial) é aquele que apresentar maior norma dentre todos os vetores de X .

Passo 2. Calcule as distâncias entre cada vetor do conjunto de treino remanescente, $(\mathbf{x}_m; m = 1, 2, \dots, M - 1)$, e o vetor-código \mathbf{w}_1^0 . O segundo vetor-código (\mathbf{w}_2^0) de W_0 é aquele que apresentar a maior distância. Obtém-se, então, um dicionário de tamanho 2, ou seja, $N = 2$.

Passo 3. Generalizando para um dicionário de tamanho $N > 2$, para cada vetor do conjunto de treino remanescente $(\mathbf{x}_m; m = 1, 2, \dots, M - N)$ calcule e armazene a menor distância entre \mathbf{x}_m e todos os N vetores do dicionário inicial obtidos até então. O vetor de treino que apresentar a maior distância ao dicionário é selecionado como o $(N + 1)$ -ésimo vetor-código de W_0 . O procedimento continua até se obter o tamanho N desejado.

3. Sequências de Baixa Discrepância

Em diversos problemas numéricos, é necessário avaliar uma dada função em todos os pontos de um dado espaço. Como isto, em geral, não é viável, é necessário estimar o valor da função mediante sua avaliação em um subconjunto, geralmente um *grid* regular, de pontos do espaço. Na escolha do subconjunto, quando não é conhecida previamente a precisão do *grid* no qual os pontos serão amostrados para obter uma resposta acurada, uma aproximação usual consiste em escolher aleatoriamente os pontos amostrais do espaço e estudar a convergência da estimativa do valor da função com base na quantidade de pontos amostrados. Neste contexto, estudos recentes [2, 16] têm demonstrado que o uso de amostragem por sequências pseudo-aleatórias, notadamente sequências de baixa discrepância, são uma alternativa com bom compromisso entre precisão e complexidade adicional de cálculo para obter estimativas precisas e com rápida convergência para avaliação de funções em diversas situações-problema.

Uma sequência de baixa discrepância é, a grosso modo, uma sequência de n -uplas que preenche uniformemente um subespaço de \mathbb{R}^n . Diferentemente das sequências

pseudo-aleatórias uniformes convencionais, tal cobertura mantém esta propriedade de uniformidade em todas as etapas da construção da sequência.

Uma definição formal de sequência de baixa discrepância segue. Sejam N_p inteiro e $E \subseteq [0, 1]$. Definimos a função de contagem $A(E; N_p) = |\{1 \leq n \leq N_p : x_n \in E\}|$, i.e, a quantidade, dentre os primeiros N_p termos da sequência, que são elementos de E . Uma sequência $\{x_n\}_{n=1}^{\infty} \subseteq I$ é definida como *uniformemente distribuída* (no intervalo $[0, 1]$) se

$$\forall a, b \in \mathbb{R}, 0 \leq a < b \leq 1 \text{ temos } \lim_{N_p \rightarrow \infty} \frac{A([a, b]; N_p)}{N_p} = b - a.$$

Tal definição é facilmente estendida para sequências multidimensionais ([11] pp 47).

Dada uma sequência finita $\{x_1, x_2, \dots, x_{N_p}\}$, definimos sua discrepância como

$$D_{N_p} = \sup_{0 \leq a < b \leq 1} \left| \frac{A([a, b]; N_p)}{N_p} - (b - a) \right|.$$

Uma sequência infinita $\{x_n\}_{n=1}^{\infty} \subseteq I$ é de baixa discrepância quando $\lim_{N_p \rightarrow \infty} D_{N_p} = 0$, i.e, uma sequência é de baixa discrepância quando tende a uma sequência uniformemente distribuída. Assim como na definição de uniformidade, a definição de discrepância é imediatamente generalizada para sequências multidimensionais ([11] pp 92 e seguintes).

Dentre as diversas sequências conhecidas de baixa discrepância, a de Van der Corput [17] e sua generalização para o caso multidimensional, dada por Halton [7], destaca-se por sua facilidade computacional de geração. Para tal sequência, vale o resultado $D_{N_p} \leq \frac{\log(N_p+1)}{N_p \log(2)}$, o que torna sua discrepância uma das de mais rápida convergência conhecidas. Formalmente, definimos a sequência de Van der Corput-Halton de dimensão K como a sequência formada pelos pontos $(\phi_{m_1}(n), \phi_{m_2}(n), \dots, \phi_{m_K}(n))$, $n = 1, 2, \dots$, em que m_1, m_2, \dots, m_K são relativamente primos dois-a-dois e $\phi_m(n) = \sum_{j=0}^s a_j m^{-j-1}$, com $n - 1 = \sum_{j=0}^s a_j m^j$ a expansão de $n - 1$ na base m_j . A K -upla m_1, m_2, \dots, m_K é dita base da sequência de Van der Corput-Halton (VDH). Se os números m_i são tomados como a sequência dos números primos, i.e., $m_1 = 2, m_2 = 3, m_3 = 5, \dots$, a sequência obtida é chamada de sequência clássica de Van der Corput-Halton.

Diferentes escolhas de bases para sequência VDH implicam coberturas diferentes do espaço $I^K = [0, 1]^K$ (i.e, o K -cubo), conforme podemos observar na Figura 1.

3.1. Aplicação na inicialização do algoritmo LBG

Seja um conjunto de treino $X = \{\mathbf{x}_m; m = 1, 2, \dots, M\}$ (ou seja, um conjunto com M vetores de treino k -dimensionais). A partir da escolha de uma base, obtém-se, da sequência de 1 até M , uma sequência de baixa discrepância com N valores, os quais determinam, dentre os M vetores k -dimensionais, aqueles que constituem o dicionário inicial do algoritmo LBG.

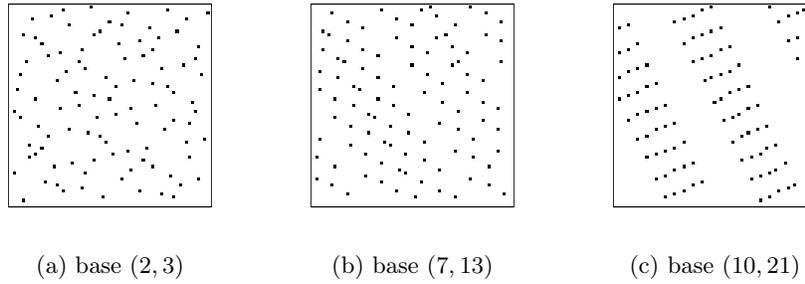


Figura 1: Coberturas distintas da região quadrada de dimensão 64×64 por 100 pontos de baixa discrepância obtidos por diferentes escolhas das bases na sequência VDH. À esquerda, base (2, 3); no centro, base (7, 13); à direita, base (10, 21). Note que a última base escolhida não é constituída de números primos, e a consequência de tal escolha é uma cobertura não uniforme da região quadrada.

4. Resultados

As imagens utilizadas nos experimentos são Barbara, Boat, Bolhas, Clock, Elaine, Goldhill, Lena, Mandrill, Peppers, Textura e Tiffany, codificadas com 256 níveis de cinza (8 bits por *pixel*), de tamanho 512×512 *pixels*, como também concatenações de 4 imagens de tamanho 256×256 *pixels*, que são: CPMB (Clock, Peppers, Mandrill e Barbara), GBTE (Goldhill, Boat, Tiffany e Elaine), MCPL (Mandrill, Clock, Peppers e Lena). As imagens estão apresentadas na Figura 2.

A taxa de codificação da QV, expressa em bits por pixel (bpp), é dada por $R = \frac{1}{k} \log_2 N$, em que N é o tamanho do dicionário (ou número de vetores de reconstrução, número de vetores-código). Neste trabalho, foram utilizados blocos de 4×4 pixels (QV com dimensão $k = 16$) e tamanhos de dicionário $N = 32, 64, 128, 256$ e 512 . Portanto, as taxas de codificação consideradas variaram de 0,3125 bpp a 0,5625 bpp. Dessa forma, para $k = 16$ e $N = 256$, tem-se uma taxa de codificação de 0,5 bpp, correspondente a uma taxa de compressão 16:1.

A avaliação de desempenho das técnicas consideradas (sequências de baixa discrepância, algoritmo de distância máxima, *splitting* e PNN) foi realizada com base na relação sinal-ruído de pico (PSNR, *peak signal to noise ratio*) das imagens reconstruídas, definida como [5]

$$\text{PSNR (dB)} = 10 \log_{10} \left[\frac{(V_p)^2}{\text{MSE}} \right], \quad (4.1)$$

em que MSE (*mean square error*) denota o erro médio quadrático entre as imagens original e reconstruída (quantizada) e V_p denota o maior valor de nível de cinza (para o caso de imagens originais 8,0 bpp, tem-se $V_p = 255$).

Nas Tabelas 1 a 5 são apresentadas as diferenças entre os valores de PSNR (dB) das imagens reconstruídas com dicionários LBG inicializados com a técnica proposta (sequência de baixa discrepância, para diversas bases) e os correspondentes valores

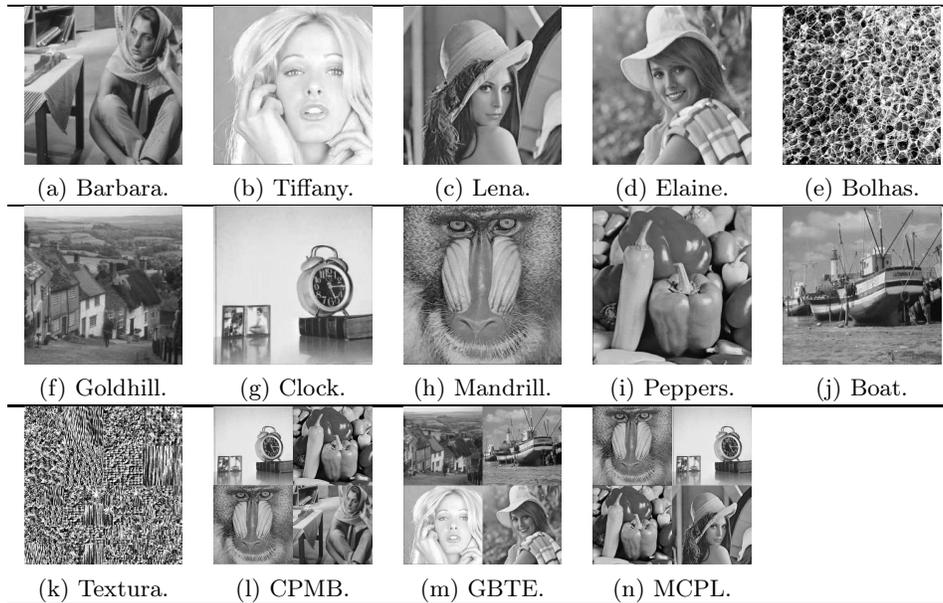


Figura 2: Imagens 512×512 usadas nas simulações.

(fixada a imagem original, a dimensão e o tamanho do dicionário) das imagens reconstruídas com dicionários LBG inicializados com a técnica de Katsavounidis et al. Nas tabelas são também apresentadas diferenças de PSNR entre *splitting* e Katsavounidis et al., bem como entre PNN e Katsavounidis et al. Como exemplo, na Tabela 1, considerando a Imagem Barbara, observa-se que um ganho de 0,33 dB em termos de PSNR da imagem reconstruída é obtido com uso de inicialização por meio de sequência de baixa discrepância de base 2 em substituição à inicialização por meio do algoritmo de distância máxima. A mesma tabela mostra que, para a imagem Clock, o uso de dicionários obtidos com LBG inicializado via sequência de baixa discrepância de base 3 leva a uma imagem reconstruída com PSNR 1,11 dB superior ao PSNR da imagem reconstruída com uso de dicionários obtidos com LBG inicializado via algoritmo de distância máxima. Valores negativos na tabela para diversas bases das sequências de baixa discrepância indicam que a inicialização via algoritmo de distância máxima leva à imagem reconstruída com PSNR superior ao obtido com uso de dicionários inicializados por sequência de baixa discrepância.

Observando a Tabela 1, constata-se que a técnica de melhor desempenho é o algoritmo PNN, o qual levou aos maiores valores de PSNR das imagens reconstruídas para 10 das 14 imagens avaliadas. O algoritmo *splitting*, por sua vez, teve o melhor desempenho para 1 das 14 imagens consideradas. Dentre as sequências de baixa discrepância, a de base 97 constituiu a melhor alternativa de inicialização. A base 97 levou aos melhores resultados para 2 das 14 imagens consideradas. Para Goldhill, por exemplo, a base 97 levou a uma imagem reconstruída com PSNR 0,49 dB acima do obtido com uso de inicialização por algoritmo de distância máxima. Para

Goldhill, o segundo melhor desempenho foi obtido como o algoritmo PNN, que levou a uma imagem reconstruída com PSNR 0,28 dB acima do obtido com uso de inicialização por algoritmo de distância máxima.

É possível observar, nas Tabelas 2, 3, 4 e 5, referentes respectivamente ao projeto de dicionários de tamanho 64, 128, 256 e 512, que o melhor desempenho, para todas as 14 imagens consideradas, está associado ao algoritmo PNN. Um aspecto deve ser destacado. No que diz respeito à complexidade computacional, avaliada por meio do número de operações lógicas e aritméticas gastas, as sequências de baixa discrepância apresentam uma superioridade sobre as demais técnicas avaliadas: correspondem à alternativa que requer o menor número de operações. O algoritmo PNN constitui o método que realiza o maior número de operações.

Tabela 1: *Diferença entre os valores de PSNR (dB) para várias bases das sequências de baixa discrepância e os correspondentes valores de PSNR (dB) obtidos com inicialização por distância máxima (algoritmo de Katsavounidis et al.), para taxa de codificação $R=0,3125$ bpp (obtida com dicionários de tamanho $N = 32$). Diferenças de valores de PSNR (dB) em relação ao algoritmo de distância máxima também são apresentadas para *splitting* e PNN.*

IMAGEM	base2	base3	base5	base7	base97	<i>Splitting</i>	PNN
Barbara	0,33	0,12	0,19	0,40	0,16	0,09	0,44
Boat	0,07	0,24	0,06	0,14	0,11	-0,04	0,25
Clock	0,75	1,11	0,91	0,78	0,90	0,65	0,58
Elaine	0,14	0,03	0,11	0,12	0,16	0,24	0,29
Goldhill	-1,03	-0,26	-0,01	-0,07	0,49	0,15	0,28
Lena	-0,63	-0,15	-0,14	-0,23	-0,20	-0,10	0,18
Mandrill	0,05	0,05	0,03	0,03	0,08	0,08	0,12
Peppers	-0,15	0,21	0,19	0,21	0,36	0,04	0,54
Tiffany	-0,58	-0,02	0,07	0,14	0,19	0,26	0,24
CPMB	0,09	0,06	0,14	0,03	-0,13	0,16	0,17
GBTE	0,06	0,00	-0,03	0,01	0,33	0,08	0,30
MCPL	0,03	0,15	0,05	0,05	0,06	0,01	0,20
Bubbles	-0,04	0,01	-0,01	0,03	-0,01	0,04	0,06
Texture	-0,01	-0,02	0,01	0,02	0,01	0,00	0,05

No que diz respeito aos resultados obtidos para tamanho de dicionário 64, conforme se pode observar na Tabela 2, o algoritmo *splitting* se apresenta como a segunda melhor opção para 7 das 14 imagens consideradas. Uma avaliação comparativa das bases revela que as bases 3 e 5 apresentam-se como as melhores opções para inicialização.

Considerando taxa de codificação 0,4375 bpp (ou seja, quantização vetorial com $k = 16$ e $N = 128$), observa-se, na Tabela 3, que a sequência de baixa discrepância de base 97 é a segunda melhor alternativa para inicialização do algoritmo LBG para 5 das 14 imagens consideradas. Para a imagem Boat, por exemplo, um ganho de 0,59 dB de PSNR da imagem reconstruída é obtido ao utilizar essa inicialização em substituição à inicialização via algoritmo de distância máxima.

Tabela 2: *Diferença entre os valores de PSNR (dB) para várias bases das sequências de baixa discrepância e os correspondentes valores de PSNR (dB) obtidos com inicialização por distância máxima (algoritmo de Katsavounidis et al.), para taxa de codificação $R=0,375$ bpp (obtida com dicionários de tamanho $N = 64$). Diferenças de valores de PSNR (dB) em relação ao algoritmo de distância máxima também são apresentadas para *splitting* e PNN.*

IMAGEM	base2	base3	base5	base7	base97	<i>Splitting</i>	PNN
Barbara	0,14	0,19	0,18	0,24	0,20	0,32	0,47
Boat	0,32	0,51	0,49	0,36	0,49	0,40	0,76
Clock	0,63	0,73	0,43	0,48	0,61	0,73	1,26
Elaine	-0,04	0,02	0,08	0,09	0,02	0,18	0,24
Goldhill	-1,19	-0,05	-0,29	-0,15	0,02	0,11	0,43
Lena	-0,76	0,02	0,05	-0,05	-0,21	-0,01	0,22
Mandrill	0,04	0,08	0,05	0,07	0,04	0,10	0,18
Peppers	-0,80	-0,06	-0,04	0,20	0,16	0,13	0,52
Tiffany	-1,51	0,30	0,44	0,36	0,42	0,42	0,50
CPMB	0,13	0,08	0,18	0,10	0,11	0,13	0,23
GBTE	0,14	-0,05	0,03	0,05	0,21	0,09	0,39
MCPL	0,08	0,15	0,08	0,14	0,14	0,05	0,36
Bubbles	-0,05	0,00	0,03	0,01	0,02	0,04	0,09
Texture	0,06	0,01	0,02	0,00	0,04	0,07	0,12

Tabela 3: *Diferença entre os valores de PSNR (dB) para várias bases das sequências de baixa discrepância e os correspondentes valores de PSNR (dB) obtidos com inicialização por distância máxima (algoritmo de Katsavounidis et al.), para taxa de codificação $R=0,4375$ bpp (obtida com dicionários de tamanho $N = 128$). Diferenças de valores de PSNR (dB) em relação ao algoritmo de distância máxima também são apresentadas para *splitting* e PNN.*

IMAGEM	base2	base3	base5	base7	base97	<i>Splitting</i>	PNN
Barbara	0,11	0,27	0,17	0,07	0,30	0,24	0,64
Boat	0,34	0,37	0,49	0,37	0,59	0,33	0,85
Clock	0,94	0,71	0,70	0,18	0,89	0,86	1,66
Elaine	0,04	0,24	0,23	0,22	0,23	0,24	0,35
Goldhill	-0,85	0,37	-0,03	0,25	0,46	0,44	1,07
Lena	-0,44	0,30	0,27	0,22	0,07	0,24	0,59
Mandrill	0,19	0,20	0,23	0,25	0,20	0,23	0,38
Peppers	-0,97	-0,14	0,09	0,22	0,11	-0,01	0,75
Tiffany	-1,45	0,27	0,32	0,40	0,36	0,32	0,61
CPMB	0,21	0,16	0,11	0,20	0,21	0,21	0,30
GBTE	0,27	0,09	0,26	0,15	0,36	0,16	0,64
MCPL	0,17	0,29	0,18	0,28	0,28	0,12	0,65
Bubbles	-0,02	0,02	-0,01	0,00	-0,01	0,02	0,16
Texture	0,07	0,04	0,05	0,06	0,03	0,07	0,18

Na Tabela 4 são apresentados resultados para dicionário de tamanho 256. Observa-se que o algoritmo *splitting* é a segunda melhor técnica de inicialização para 5 das 14 imagens consideradas. Uma avaliação comparativa das bases revela que a melhor inicialização está associada à base 7. De fato, trata-se da segunda melhor alternativa de inicialização para 3 das 14 imagens consideradas, a saber: Elaine, Tiffany e CPMB. Para esta última, um ganho de 0,36 dB de PSNR da imagem reconstruída é obtido ao se utilizar dicionário obtido com o algoritmo LBG inicializado com sequência de baixa discrepância de base 7 em substituição ao dicionário obtido com o algoritmo LBG inicializado via algoritmo de distância máxima.

Tabela 4: Diferença entre os valores de PSNR (dB) para várias bases das sequências de baixa discrepância e os correspondentes valores de PSNR (dB) obtidos com inicialização por distância máxima (algoritmo de Katsavounidis et al.), para taxa de codificação $R=0,5$ bpp (obtida com dicionários de tamanho $N = 256$). Diferenças de valores de PSNR (dB) em relação ao algoritmo de distância máxima também são apresentadas para *splitting* e PNN.

IMAGEM	base2	base3	base5	base7	base97	<i>Splitting</i>	PNN
Barbara	0,15	0,26	0,16	0,30	0,22	0,35	0,90
Boat	0,28	0,49	0,50	0,41	0,65	0,45	1,33
Clock	-0,41	0,26	0,14	0,28	0,33	0,96	1,55
Elaine	0,16	0,19	0,16	0,20	0,18	0,16	0,44
Goldhill	-0,50	0,28	0,02	0,36	0,31	0,45	1,12
Lena	0,02	0,09	0,11	0,09	-0,01	0,14	0,69
Mandrill	0,26	0,25	0,28	0,22	0,25	0,18	0,55
Peppers	-0,59	0,02	0,04	-0,03	-0,10	0,09	1,11
Tiffany	-0,40	0,32	0,33	0,34	0,25	0,31	0,77
CPMB	0,29	0,27	0,25	0,36	0,28	0,35	0,60
GBTE	0,15	0,11	0,20	0,14	0,16	0,06	0,78
MCPL	0,07	0,31	0,28	0,29	0,27	0,18	0,87
Bubbles	0,13	0,09	0,08	0,11	0,09	0,11	0,36
Texture	0,13	0,11	0,12	0,10	0,11	0,08	0,36

No tocante aos resultados obtidos para dicionário de tamanho 512, o exame da Tabela 5 revela que a base 97 constitui a segunda melhor alternativa de inicialização para 3 das 14 imagens consideradas, a saber: Boat, GBTE e texture. O algoritmo *splitting* é a segunda melhor alternativa de inicialização para 7 imagens.

A Figura 3 apresenta diferenças entre os valores de PSNR das imagens Boat reconstruídas com uso de dicionários LBG obtidos via inicialização por sequência de baixa discrepância e os correspondentes valores de PSNR das imagens reconstruídas com dicionários LBG obtidos via inicialização por algoritmo de distância máxima. A figura também apresenta diferenças entre valores de PSNR obtidos com *splitting* e os correspondentes valores obtidos com o algoritmo de distância máxima. Diferenças entre valores de PSNR obtidos com o algoritmo PNN e os correspondentes valores obtidos com o algoritmo de distância máxima também são apresentados na figura. Em geral, observa-se que o melhor desempenho é apresentado pelo algoritmo PNN. De fato, para todos os valores de taxa de codificação considerados, o algoritmo

Tabela 5: Diferença entre os valores de PSNR (dB) para várias bases das sequências de baixa discrepância e os correspondentes valores de PSNR (dB) obtidos com inicialização por distância máxima (algoritmo de Katsavounidis et al.), para taxa de codificação $R=0,5625$ bpp (obtida com dicionários de tamanho $N = 512$). Diferenças de valores de PSNR (dB) em relação ao algoritmo de distância máxima também são apresentadas para *splitting* e PNN.

IMAGEM	base2	base3	base5	base7	base97	<i>Splitting</i>	PNN
Barbara	0,07	0,07	0,22	0,21	0,21	0,22	1,24
Boat	0,45	0,52	0,56	0,53	0,68	0,66	1,97
Clock	0,66	0,52	0,42	0,38	0,66	1,18	2,18
Elaine	0,12	0,16	0,17	0,17	0,18	0,22	0,65
Goldhill	-0,07	0,28	0,37	0,30	0,40	0,45	1,39
Lena	-0,19	-0,03	-0,01	0,09	0,03	0,23	0,87
Mandrill	0,35	0,34	0,30	0,34	0,32	0,22	0,90
Peppers	-0,26	0,02	-0,05	0,09	-0,04	0,23	1,85
Tiffany	-0,06	0,28	0,30	0,29	0,24	-2,77	1,12
CPMB	0,01	0,02	0,22	-0,02	0,18	0,39	0,58
GBTE	0,11	0,06	0,10	0,08	0,19	0,03	1,21
MCPL	0,26	0,32	0,38	0,32	0,34	0,19	1,40
Bubbles	0,12	0,16	0,13	0,11	0,13	0,08	0,69
Texture	0,11	0,17	0,17	0,19	0,21	0,14	0,66

PNN leva aos maiores valores de PSNR das imagens reconstruídas. Dentre as bases avaliadas, a base 97 constitui-se como a melhor opção para inicialização. Dentre as bases avaliadas, os valores mais baixos de PSNR das imagens reconstruídas foram obtidos com a base 2. Vale salientar, no entanto, que, para todas as taxas de codificação consideradas, os resultados obtidos com sequência de baixa discrepância de base 2 foram melhores que os obtidos pelo algoritmo de distância máxima.

Para a imagem GBTE, conforme mostra a Figura 4, o algoritmo PNN constitui a técnica de melhor desempenho. Uma avaliação comparativa das bases avaliadas revela que a base 97 é a que leva aos melhores resultados para a maioria das taxas de codificação consideradas: uma exceção é a taxa 0,5 bpp, para a qual o melhor desempenho está associado à base 5; para a imagem GBTE, a base 3 foi a que levou aos piores resultados, um dos quais inferior ao obtido com o algoritmo de distância máxima, conforme indica o valor negativo da curva para a base 3, para taxa de 0,375 bpp. Observa-se, na Figura 4, que ganhos de até 0,35 dB podem ser obtidos ao se substituir o algoritmo de distância máxima pela sequência de baixa discrepância de base 97 como alternativa de inicialização do algoritmo LBG.

5. Conclusão

Neste trabalho foi proposto o uso de sequências de baixa discrepância como alternativa de inicialização do algoritmo LBG, o qual constitui a técnica mais amplamente utilizada para o projeto de quantizadores vetoriais. Os resultados obtidos com as

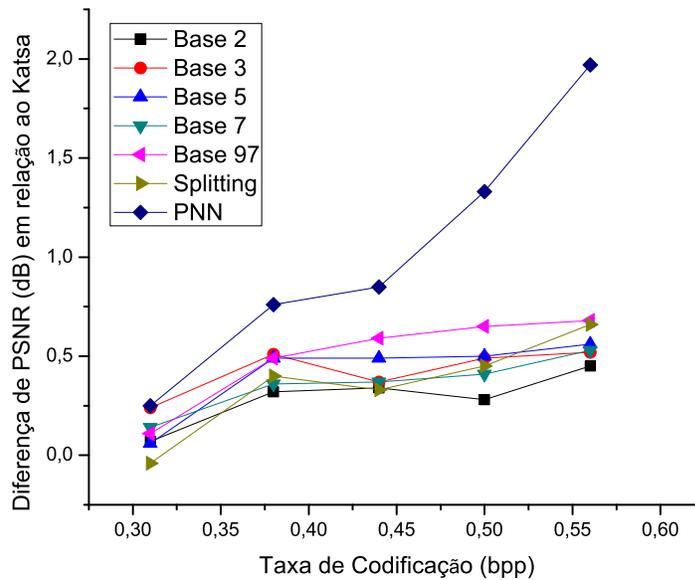


Figura 3: Diferença de PSNR (dB) das imagens Boat reconstruídas, em relação ao algoritmo de distância máxima, em função da taxa de codificação, para diversas escolhas de base.

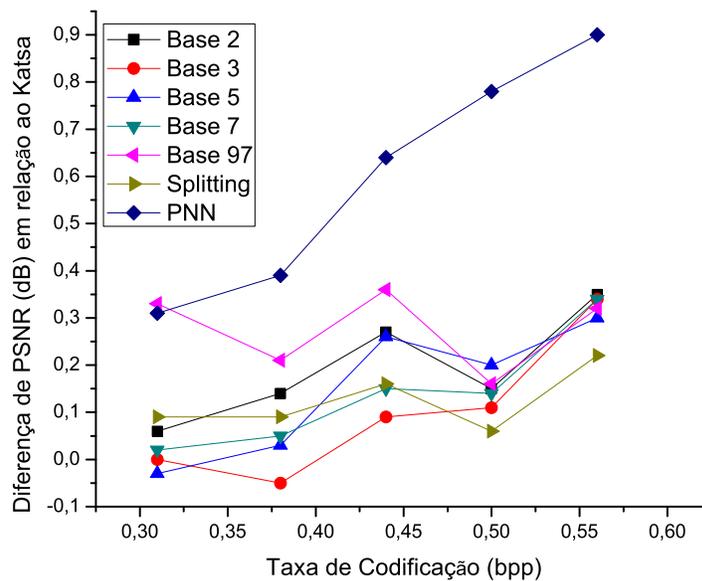


Figura 4: Diferença de PSNR (dB) das imagens GBTE reconstruídas, em relação ao algoritmo de distância máxima, em função da taxa de codificação, para diversas escolhas de base.

seqüências de baixa discrepância foram comparados ao algoritmo de distância máxima. Outras duas técnicas foram consideradas para o propósito de comparação de desempenho: os algoritmos *splitting* e *pairwise nearest neighbor* (PNN). Em compressão de imagens baseada em quantização vetorial, foram avaliadas seqüências de baixa discrepância de bases 2, 3, 5, 7 e 97. Os resultados obtidos, para taxas de codificação compreendidas entre 0,3125 bpp e 0,5625 bpp, revelaram que imagens reconstruídas com melhor qualidade (ou seja, com maiores valores de relação sinal-ruído de pico) são obtidas ao se utilizar o algoritmo PNN. Este método, entretanto, é o que realiza o maior número de operações lógicas e aritméticas, dentre os considerados neste trabalho. Observou-se que a inicialização por seqüência de baixa discrepância tem um desempenho superior ao obtido com a inicialização por algoritmo de distância máxima. Para um conjunto de 14 imagens avaliadas, dentre as bases consideradas, a base 97 pode, em geral, ser apontada como a melhor para o propósito de inicialização do algoritmo LBG. Para a codificação da imagem Boat à taxa de 0,5625 bpp, por exemplo, a substituição do algoritmo de distância máxima pela seqüência de baixa discrepância de base 97 leva a um ganho de aproximadamente 0,7 dB em termos de PSNR da imagem reconstruída.

Em trabalhos futuros, será avaliado se é possível definir uma escolha de base em função de características da imagem utilizada como conjunto de treino.

Abstract. Speech and image compression systems, steganography and pattern recognition are examples of applications of vector quantization (VQ), whose performance depends strongly on the quality of the designed codebooks. A widely used technique for VQ codebook design is the LBG (Linde-Buzo-Gray) algorithm. In the present paper, low discrepancy sequences are presented as an alternative for the LBG initialization. The performance of the proposed technique is assessed in the scenario of image compression based on VQ.

Referências

- [1] Y.K. Chiang, P. Tsai, Steganography using overlapping codebook partition, *Signal Processing*, **88**, No. 5 (2008), 1203–1215.
- [2] I. Dalal, D. Stefan, J. Harwayne-Gidansky, Low discrepancy sequences for Monte Carlo simulations on reconfigurable platforms, In: “Proceedings of the International Conference on Application-Specific Systems, Architectures and Processors, 2008”, pp. 108–113, 2008.
- [3] W.H. Equitz, A new vector quantization clustering algorithm, *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**, No. 10 (1989), 1568–1575.
- [4] P. Fränti, Genetic algorithm with deterministic crossover for vector quantization, *Pattern Recognition Letters*, **21**, No. 1 (2000), 61–68.
- [5] A. Gersho, R.M. Gray, “Vector Quantization and Signal Compression”, Kluwer Academic Publishers, Boston, 1992.
- [6] R.M. Gray, Vector quantization, *IEEE ASSP Magazine*, (1984), 4–29.

- [7] J.H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals, *Numer. Math.*, (1960), 84–90.
- [8] Y. Hu, High-capacity image hiding scheme based on vector quantization, *Pattern Recognition*, **3**, (2006), 1715–1724.
- [9] I. Katsavounidis, C.-C.J. Kuo, Z. Zhang, A new initialization technique for generalized lloyd iteration, *IEEE Signal Processing Letters*, **1**, No. 10, (1994), 144–146.
- [10] A. Kjoelen, S.E. Umbaugh, M. Zuke, Compression of skin tumor images – wavelet/vector quantization methods for reducing the time, cost and bandwidth of storing and transmitting data, *IEEE Engineering in Medicine and Biology*, (1998), 73–80.
- [11] N.H. Kuipers, “Uniform Distribution of Sequences”, Addison Wesley Publishing Company, Inc., New York, 1992.
- [12] Y. Linde, A. Buzo, R.M. Gray, An algorithm for vector quantizer design, *IEEE Transactions on Communications*, **28**, No. 1 (1980), 84–95.
- [13] W.T.A. Lopes, F. Madeiro, B.G. Aguiar Neto, M.S. Alencar, Combining modulation diversity and index assignment to improve image VQ for a Rayleigh fading channel, *Learning and Nonlinear Models – Revista da Sociedade Brasileira de Redes Neurais*, **1**, No. 3 (2004), 160–171.
- [14] F. Madeiro, J.M. Fachine, W.T.A. Lopes, B.G. Aguiar Neto, M.S. Alencar, Identificação Vocal por Frequência Fundamental, QV e HMMS, In: “EMTOM-Ação: A Prosódia em Perspectiva”, (M.A.M. Aguiar and F. Madeiro, eds.), chapter 4, pp. 91–120. Editora Universitária da UFPE, Recife, PE, first edition, 2007.
- [15] K.K. Paliwal, B.S. Atal, Efficient vector quantization of LPC parameters at 24 bits/frame, *IEEE Transactions on Speech and Audio Processing*, **1**, No. 1 (1993), 3–14.
- [16] O. Teytaud, When does quasi-random work?, In: “Proceedings of the 10th International Conference on Parallel Problem Solving from Nature: PPSN X”, pp. 325–336, 2008.
- [17] Van der Corput, Zahlentheoretische abschaätzungen, *Math. Ann.*, (1921), 53–79.
- [18] P. Yahampath, P. Rondeau, Multiple-description predictive-vector quantization with applications to low bit-rate speech coding over networks, *IEEE Transactions on Audio, Speech, and Language Processing*, **15**, No. 3 (2007), 749–755.