

AVALIAÇÃO DO DESEMPENHO DE ALGORITMOS DE COMPRESSÃO DE
IMAGENS USANDO RECORRÊNCIA DE PADRÕES MULTIESCALAS

Frederico da Silva Pinagé

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS
EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof. Eduardo Antônio Barros da Silva, Ph.D.

Prof. Gelson Vieira Mendonça, Ph.D.

Prof. Abraham Alcaim, Ph.D.

Prof. Weiler Alves Finamore, Ph.D.

Prof. Murilo Bresciani de Carvalho, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

JULHO DE 2005

PINAGÉ, FREDERICO DA SILVA

Avaliação do desempenho de algoritmos
de compressão de imagens usando
recorrência de padrões multiescalas
[Rio de Janeiro] 2005

XVI,97 pp 29,7 cm (COPPE/UFRJ, M.Sc.,
Engenharia Elétrica, 2005)

Tese - Universidade Federal do Rio de
Janeiro, COPPE

- 1.Compressão de Imagens
- 2.Casamento de Padrões Multiescalas
- 3.Otimização Taxa-distorção
- 4.Levantamento Estatístico
- 5.Avaliação do desempenho

I.COPPE/UFRJ II.Título (série)

Agradecimentos

Primeiramente, o meu reconhecimento aos meus pais Frederico e Neide Pinagé que acompanharam toda a minha vida estudantil e atualmente, profissional e como mestrando, às minhas irmãs Kellen e Gizele Pinagé, pelo incentivo aos estudos e desenvolvimento do trabalho do mestrado.

Agradeço às pessoas que muito me ajudaram a escrever este trabalho, por suas reminiscências, paciência e orientação, sou grato ao Eddie Lima, Lúcio Cauper, Ricardo Cruz, Úrsula Abecassis, aos alunos da FACUPI que também são meu grande incentivo na vida acadêmica, e a legião de amigos e colegas.

Agradeço especialmente ao Eduardo Barros (COPPE), meu orientador, pela sua paciência e por ter conduzido com equilíbrio necessário esse projeto, e também ao Professor Gelson (COPPE). E aos professores Cícero e Marly Costa, pelas suas orientações em iniciação científica e por me mostrarem o mundo da pesquisa.

Agradeço também a FUCAPI, por permitir a realização deste trabalho, e em especial a Niomar Pimenta e Antonio Luiz Maués, que me apoiaram e incentivaram no mestrado e o Waldir Sabino, que me ajudou nas explicações, incentivos e conselhos.

E como não poderia deixar de ser, meus agradecimentos à equipe de técnicos da SUFRAMA e UFAM que tornaram possível a realização do curso de Mestrado em Manaus.

Frederico da Silva Pinagé

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

AVALIAÇÃO DO DESEMPENHO DE ALGORITMOS DE COMPRESSÃO DE IMAGENS USANDO RECORRÊNCIA DE PADRÕES MULTIESCALAS

Frederico da Silva Pinagé

Julho/2005

Orientadores: Eduardo Antônio Barros da Silva
Gelson Vieira Mendonça

Programa: Engenharia Elétrica

O método de compressão de sinais MMP (*“Multi-dimensional Multiscale Parser”*) é baseado em recorrência de padrões multiescala onde cada bloco do sinal de entrada é aproximado por um dicionário adaptativo sendo atualizado com versões dilatadas e contraídas de concatenações dos blocos previamente codificados. Neste trabalho investigamos, através do levantamento estatístico das palavras do dicionário, variações do MMP entre as quais a atualização do dicionário com a inclusão de blocos rotacionados e transpostos, bem como de blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho. Além disso, investiga-se uma modificação na estrutura do dicionário, que consiste em indexar suas palavras dependendo de a partir de qual escala elas foram introduzidas no mesmo (segmentação no espaço de códigos). Os resultados em relação ao MMP, para imagens suaves, corresponderam a um ganho de aproximadamente 0.5dB @ 0.5bpp para a imagem Lena e ganhos de aproximadamente 0.3dB @ 0.5bpp para imagens combinando texto e gráficos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

EVOLUTION OF THE PERFORMANCE OF ALGORITHMS OF
COMPRESSION OF IMAGES USING RECURRENT MULTISCALE
PATTERNS

Frederico da Silva Pinagé

July/2005

Advisors: Eduardo Antônio Barros da Silva

Gelson Vieira Mendonça

Department: Electrical Engineering

The MMP (Multi-dimensional Multiscale Parser) signal compression method is based on multiscale recurrent patterns, where each input block is approximated by an adaptive dictionary which is updated with dilated and contracted versions of previously coded blocks. In this work we investigate MMP variations through statistic survey of update dictionary word usage, the variations include dictionary with rotated and transposed blocks as well as displaced blocks of multiples of $1/2$, $1/4$ and $1/8$ of their size. In addition, we investigate a modification in the structure of the dictionary that consist in indexing their words depending on from which scale they were introduced in the dictionary (code-space segmentation). The results in relation to MMP for soft images have shows gains of approximately 0.5dB @ 0.5bpp for the image Lena and of approximately 0.3dB @ 0.5bpp for images that combine text and graphics.

Sumário

1	Introdução	1
2	Compressão de dados	3
2.1	Técnicas de compressão de dados	3
2.1.1	Compressão sem perdas	4
2.1.2	Compressão com perdas	5
2.2	Teoria da informação para compressão sem perdas	8
2.3	Teoria da informação para compressão com perdas	9
2.4	Medidas de desempenho	10
3	O MMP	12
3.1	Descrição do MMP	12
3.2	Versões do MMP	18
3.2.1	MMP com Codificador Aritmético	19
3.2.2	MMP Otimização Taxa-distorção (MMPRD)	19
3.2.3	Otimização taxa-distorção	19
3.2.4	MMP com Otimização taxa-distorção intermediário (MMPRDI)	22
4	MMPRDI usando segmentação do dicionário do espaço de códigos	23
4.1	Motivação	24
4.2	Descrição	25
4.3	Atualização do dicionário	26
4.4	Resultados	26
4.4.1	Curvas taxa-distorção	34

5	MMPRDI usando blocos rotacionados e transpostos	39
5.1	Descrição	39
5.2	Atualização do Dicionário	40
5.3	Curvas taxa-distorção	40
5.4	Resultados	47
5.4.1	Resultados dos blocos rotações e transpostas mais usados . . .	52
6	MMPRDI com deslocamento de blocos	58
6.1	Descrição	58
6.2	Atualização do Dicionário	59
6.3	Curva taxa-distorção	59
6.4	Resultados	64
6.4.1	Resultados para adição de blocos deslocados de múltiplos de 1/2 de seu tamanho	64
6.4.2	Resultados de blocos deslocados mais usados com desloca- mento de múltiplos de 1/2 de seu tamanho	68
6.4.3	Resultados para adição de blocos deslocados de múltiplos de 1/4 de seu tamanho	72
6.4.4	Resultados de blocos deslocados mais usados com desloca- mento de múltiplos de 1/4 de seu tamanho	75
6.4.5	Resultados para adição de blocos deslocados de múltiplos de 1/8 de seu tamanho	79
6.4.6	Resultados de blocos deslocados mais usados com desloca- mento de múltiplos de 1/8 de seu tamanho	83
6.4.7	Levantamento Estatístico: conclusões	87
7	Conclusão	88
	Referências Bibliográficas	89
A	Imagens	93
A.1	Imagem <i>aerial</i>	93
A.2	Imagem <i>Baboon</i>	94
A.3	Imagem <i>Barbara</i>	94

A.4	Imagem <i>Bridge</i>	95
A.5	Imagem <i>f16</i>	95
A.6	Imagem <i>gold</i>	96
A.7	Imagem <i>Lena</i>	96
A.8	Imagem <i>pp1205</i>	97
A.9	Imagem <i>pp1209</i>	97

Lista de Figuras

2.1	Esquema de Técnica de Compressão de dados	4
2.2	Esquema de Técnica de Compressão sem perdas	5
2.3	Esquema de Técnica de Compressão com perdas	5
3.1	Esquema do Dicionário do MMP	13
3.2	Exemplo de um bloco retirado da Imagem para processamento	14
3.3	Projeção do Bloco de Entrada X (parte da imagem) no dicionário . .	14
3.4	Atualização do Dicionário	15
3.5	Exemplo do Bloco X segmentado	16
3.6	Árvore de segmentação do Bloco de Entrada X	16
3.7	Projeção dos segmentos dos blocos de entrada X (X^1 e X^2 no dicionário)	17
3.8	Árvore de segmentação aproximada do Bloco de Entrada X	18
3.9	Árvore de segmentação representada com nodos folhas	21
4.1	Novo esquema do dicionário para avaliação do desempenho do MMPRDI	24
4.2	Estrutura do dicionário para avaliação do desempenho do MMPRDI .	25
4.3	Projeção do bloco de entrada X no dicionário	27
4.4	Exemplo da atualização do dicionário	28
4.5	Resultado do desempenho da imagem de teste aerial (apêndice A.1). .	30
4.6	Resultado do desempenho da imagem de teste baboon (apêndice A.2). .	30
4.7	Resultado do desempenho da imagem de teste barbara (apêndice A.3). .	31
4.8	Resultado do desempenho da imagem de teste bridge (apêndice A.4). .	31
4.9	Resultado do desempenho da imagem de teste f16 (apêndice A.5). . .	32
4.10	Resultado do desempenho da imagem de teste gold (apêndice A.6). .	32
4.11	Resultado do desempenho da imagem de teste lena (apêndice A.7). .	33
4.12	Resultado do desempenho da imagem de teste pp1205 (apêndice A.8). .	33

4.13	Resultado do desempenho da imagem de teste pp1209 (apêndice A.9).	34
4.14	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem Barbara (apêndice A.3) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	36
4.15	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem F16 (apêndice A.5) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	36
4.16	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem Gold (apêndice A.6) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	37
4.17	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem Lena (apêndice A.7) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	37
4.18	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem PP1205 (apêndice A.8) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	38
4.19	Curva Taxa-distorção do MPRDI, MPRDI-seg para imagem PP1209 (apêndice A.9) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).	38
5.1	Exemplo de um bloco concatenado e suas rotações, transposta e rotações de sua transposta. (a) mostra o bloco concatenado para atualizar o dicionário e para realizar as rotações, além da transposta, e rotações do bloco transposto; (b), (c) e (d) São as rotações de 90° , 180° e 270° do bloco concatenado, respectivamente; (e) Transposta do bloco concatenado; (f), (g) e (h) Bloco transposto rotacionado de 90° , 180° e 270° , respectivamente.	41
5.2	Curva Taxa-distorção do MPRDI, e blocos rotacionados e transpostos para imagem Aerial 512×512 .	42
5.3	Curva Taxa-distorção do MPRDI, e blocos rotacionados e transpostos para imagem Baboon 512×512 .	43
5.4	Curva Taxa-distorção do MPRDI, e blocos rotacionados e transpostos para imagem Barbara 512×512 .	43

5.5	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Bridge 512×512	44
5.6	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem F16 512×512	44
5.7	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Gold 512×512	45
5.8	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Lena 512×512	45
5.9	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem PP1205 512×512	46
5.10	Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem PP1209 512×512	46
5.11	Resultado do levantamento estatístico do MMPRDI para a imagem de teste aerial com rotação e transposta de blocos.	48
5.12	Resultado do levantamento estatístico do MMPRDI para a imagem de teste baboon com rotação e transposta de blocos.	48
5.13	Resultado do levantamento estatístico do MMPRDI para a imagem de teste barbara com rotação e transposta de blocos.	49
5.14	Resultado do levantamento estatístico do MMPRDI para a imagem de teste bridge com rotação e transposta de blocos.	49
5.15	Resultado do levantamento estatístico do MMPRDI para a imagem de teste f16 com rotação e transposta de blocos.	50
5.16	Resultado do levantamento estatístico do MMPRDI para a imagem de teste gold com rotação e transposta de blocos.	50
5.17	Resultado do levantamento estatístico do MMPRDI para a imagem de teste lena com rotação e transposta de blocos.	51
5.18	Resultado do levantamento estatístico do MMPRDI para a imagem de teste pp1205 com rotação e transposta de blocos.	51
5.19	Resultado do levantamento estatístico do MMPRDI para a imagem de teste pp1209 com rotação e transposta de blocos.	52
5.20	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste aerial.	53

5.21	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste baboon.	54
5.22	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste barbara.	54
5.23	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste bridge.	55
5.24	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste f16.	55
5.25	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste gold.	56
5.26	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste lena.	56
5.27	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste pp1205.	57
5.28	Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste pp1209.	57
6.1	Exemplo de como será feito o deslocamento de blocos com comprimentos determinados. Neste exemplo o bloco utilizado tem tamanho 8×8	60
6.2	Curva Taxa-distorção do MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem Barbara 512×512 . . .	61
6.3	Curva Taxa-distorção do MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem F16 512×512 . . .	61
6.4	Curva Taxa-distorção do MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem Gold 512×512 . . .	62
6.5	Curva Taxa-distorção do MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem Lena 512×512 . . .	62
6.6	Curva Taxa-distorção do MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem PP1205 512×512 . . .	63
6.7	Curva Taxa-distorção do MMP, MMPRDI, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho para imagem PP1209 512×512	63

6.8	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara com blocos deslocados de múltiplos de 1/2 do seu tamanho. .	65
6.10	Resultado do levantamento estatístico do MMPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/2 do seu tamanho. . .	65
6.9	Resultado do levantamento estatístico do MMPRDI para a imagem F16 com blocos deslocados de múltiplos de 1/2 do seu tamanho. . . .	66
6.11	Resultado do levantamento estatístico do MMPRDI para a imagem Lena com blocos deslocados de múltiplos de 1/2 do seu tamanho. . .	66
6.12	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 com blocos deslocados de múltiplos de 1/2 do seu tamanho. .	67
6.13	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/2 do seu tamanho. .	67
6.14	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.	69
6.16	Resultado do levantamento estatístico do MMPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.	69
6.15	Resultado do levantamento estatístico do MMPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.	70
6.17	Resultado do levantamento estatístico do MMPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.	70
6.18	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.	71
6.19	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.	71
6.20	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara com blocos deslocados de múltiplo de 1/4 do seu tamanho. .	72

6.21	Resultado do levantamento estatístico do MMPRDI para a imagem F16 com blocos deslocados de múltiplo de 1/4 do seu tamanho.	73
6.22	Resultado do levantamento estatístico do MMPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/4 do seu tamanho.	73
6.23	Resultado do levantamento estatístico do MMPRDI para a imagem Lena com blocos deslocados de múltiplo 1/4 do seu tamanho.	74
6.24	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 com blocos deslocados de múltiplos 1/4 do seu tamanho.	74
6.25	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/4 do seu tamanho.	75
6.26	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	76
6.27	Resultado do levantamento estatístico do MMPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	77
6.28	Resultado do levantamento estatístico do MMPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	77
6.29	Resultado do levantamento estatístico do MMPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	78
6.30	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	78
6.31	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.	79
6.32	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara com blocos deslocados de múltiplos de 1/8 do seu tamanho.	80
6.34	Resultado do levantamento estatístico do MMPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/8 do seu tamanho.	80

6.33	Resultado do levantamento estatístico do MMPRDI para a imagem F16 com blocos deslocados de múltiplos de 1/8 do seu tamanho.	81
6.35	Resultado do levantamento estatístico do MMPRDI para a imagem Lena com blocos deslocados de múltiplos de 1/8 do seu tamanho.	81
6.36	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 com blocos deslocados de múltiplos de 1/8 do seu tamanho.	82
6.37	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/8 do seu tamanho.	82
6.38	Resultado do levantamento estatístico do MMPRDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	84
6.39	Resultado do levantamento estatístico do MMPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	84
6.40	Resultado do levantamento estatístico do MMPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	85
6.41	Resultado do levantamento estatístico do MMPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	85
6.42	Resultado do levantamento estatístico do MMPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	86
6.43	Resultado do levantamento estatístico do MMPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.	86
A.1	Imagem <i>aerial</i>	93
A.2	Imagem <i>baboon</i>	94
A.3	Imagem <i>barbara</i>	94
A.4	Imagem <i>bridge</i>	95
A.5	Imagem <i>f16</i>	95
A.6	Imagem <i>gold</i>	96

A.7	Imagem <i>lena</i>	96
A.8	Imagem <i>pp1205</i>	97
A.9	Imagem <i>pp1209</i>	97

Capítulo 1

Introdução

Nos últimos anos houve um crescimento na utilização de imagens digitais em diversas áreas, como por exemplo na transmissão de dados: video conferência, radar e sonar (operações militares), telemedicina; além disso houve uma grande evolução da padronização de transmissão de vídeo. Também podemos citar a tomografia computadorizada, ressonância magnética, além da computação multimídia e imagens de satélite. Dados contendo imagens digitais são transmitidos de um lado a outro, nos quatro cantos do planeta, e são armazenados. E para esta informação ser transmitida com mais rapidez e utilizar uma quantidade menor de memória para armazenamento, é necessário comprimí-la. Os equipamentos utilizados em vídeo conferência, em telemedicina e computação multimídia, utilizam a compressão de imagens. Devido à grande quantidade de bits que uma imagem digital necessita para ser representada, os algoritmos utilizados em compressão de imagens têm grande importância nas aplicações citadas acima. Em alguns casos esses algoritmos têm a saída idêntica à imagem original (algoritmo de compressão sem perdas) e em outros, há a diferença entre a saída do algoritmo e a imagem original (algoritmo de compressão com perdas).

No caso propomos um algoritmo de compressão com perdas baseado em casamento aproximado de padrões recorrentes multiescalas [1], isto é, um algoritmo que tenta representar um bloco de entrada (um bloco retirado da imagem) através de aproximações dilatadas ou contraídas de blocos previamente codificados.

Neste trabalho, fazemos uma análise detalhada do desempenho de codificação de esquemas de compressão de sinais baseados no MMP (“Multi-dimensional Multi-

scale Parser”), para, a partir daí, propor modificações que possam redundar em uma melhora de desempenho. As modificações propostas se relacionam essencialmente à construção e atualização do dicionário. Elas consistem em segmentar o dicionário no espaço de códigos (ou seja, indexar as palavras do dicionário dependendo de a partir de qual escala elas foram introduzidas no mesmo) e, além disso, incluir no dicionário blocos rotacionados e transpostos, e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho. A tese se divide em 7 capítulos.

O capítulo 2 apresenta uma definição de compressão de dados, os tipos de compressão de dados, introduz conceitos de teoria da informação e apresenta alguns algoritmos de compressão de dados, bem como suas medidas de desempenho.

O capítulo 3 descreve o algoritmo MMP, as suas versões e as modificações feitas em sua estrutura.

No capítulo 4 inicialmente descrevemos o MMP com a segmentação do dicionário no espaço de códigos, destacando a atualização do dicionário e a codificação de seus índices. Fazemos também um levantamento estatístico da utilização das palavras do dicionário, a fim de verificar se é efetivo ou não o uso da segmentação no espaço de códigos. Em seguida, apresentamos os resultados obtidos.

O capítulo 5 investiga outra variação do algoritmo MMP, que consiste na inclusão de blocos rotacionados e transpostos no seu dicionário. Apresentamos também os resultados do levantamento estatístico das palavras do dicionário, além das curvas taxa-distorção do algoritmo em questão.

O capítulo 6 investiga a inclusão de blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho no MMP. Mostramos também os resultados do levantamento estatístico das palavras do dicionário, além das curvas taxa-distorção do algoritmo em questão.

Nós concluímos o trabalho no capítulo 7, onde resumimos as idéias dos capítulos, comentamos os resultados e apresentamos sugestões de investigações para trabalhos futuros.

Capítulo 2

Compressão de dados

A compressão de dados consiste em reduzir o número de bits necessários para representar uma informação (por exemplo, uma imagem ou uma seqüência de vídeo). Podemos representar a informação de forma compacta identificando e usando estruturas que existem no sinal. A informação pode estar representada na forma de caracteres num arquivo texto, número de amostras de fala, as formas de onda de uma imagem, entre outros. E pelo fato de sempre estarmos gerando e usando informações na forma digital, torna-se cada vez mais necessário o uso da compressão de dados para podermos representar a informação usando um número de bits reduzidos. Existem vários tipos de estruturas que são usadas na compressão de dados. Por exemplo, seja um arquivo texto contendo alguns caracteres que ocorrem com mais freqüência que os outros. Estes caracteres podem ser representados com um menor número de bits, e os que ocorrem com menor freqüência, representá-los com um número de bits maior. Na verdade podemos nos utilizar de vários tipos de estruturas contidas num sinal para obter compressão.

Na próxima seção falaremos dos tipos de técnicas de compressão que podemos utilizar para reduzir o número de bits necessários para representar um determinado sinal.

2.1 Técnicas de compressão de dados

Quando falamos em técnicas de compressão (ou algoritmos de compressão), estamos referindo a um algoritmo de codificação (codificador) que processará uma

entrada X e produzirá uma saída X_c , e a um algoritmo de decodificação (decodificador) que reconstruirá a informação Y a partir da saída do codificador, ou seja, X_c . Esta técnica é representada pelo esquema da figura 2.1.

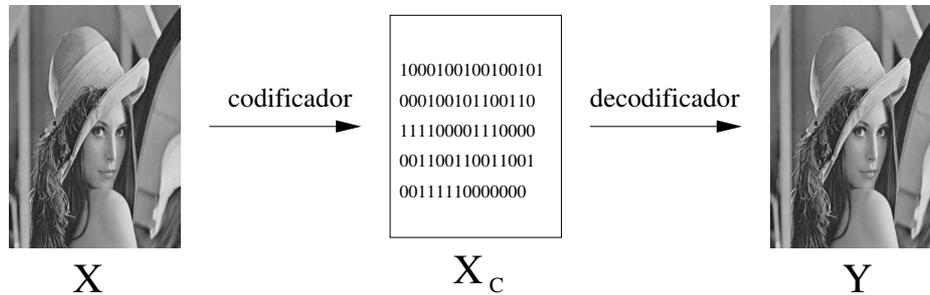


Figura 2.1: Esquema de Técnica de Compressão de dados

Com base no esquema da figura 2.1, podemos obter dois tipos de compressão de dados: compressão sem perdas (figura 2.2), na qual a saída Y do decodificador é idêntica à entrada X do codificador (não há diferença alguma entre a entrada X do codificador e a saída) e a compressão com perdas (figura 2.3), na qual a saída Y do decodificador é diferente da entrada X do codificador, ou seja, existe diferença entre a saída Y e a entrada X .

Nas seções seguintes iremos comentar sobre estes dois tipos de compressão de dados.

2.1.1 Compressão sem perdas

Como mencionamos anteriormente, nas técnicas de compressão sem perdas os dados recuperados (saída do decodificador) são idênticos aos dados presentes na entrada. Isto é, podemos recuperar completamente a informação à entrada quando efetuamos a decodificação. Uma característica destas técnicas (que pode ser encarada como uma limitação) é o fato de que não se consegue uma alta taxa de compressão.

Usamos compressão sem perdas em aplicações que exigem que não haja diferença entre a informação reconstruída e a original. Por exemplo, em imagens de radiografia, uma pequena diferença entre a imagem reconstruída e a imagem original, pode levar a um diagnóstico errado. As técnicas de compressão sem perdas também têm aplicação em compressão de textos e de imagens de satélite.

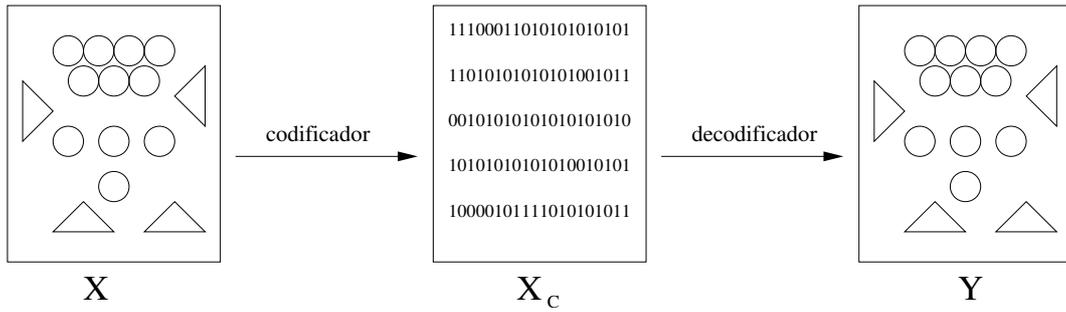


Figura 2.2: Esquema de Técnica de Compressão sem perdas

Existem várias técnicas que utilizam este tipo de compressão sem perdas, dentre elas podemos citar a técnica desenvolvida por David Huffman, chamada código de Huffman [2, 3], codificação aritmética através de Peter Elias e descrita em [4], e as técnicas baseadas em dicionários, que foram desenvolvidas por Abraham Lempel e Jacob Ziv [5, 6].

Há muitas situações que exigem que os dados recuperados sejam idênticos à entrada, como em alguns exemplos citados anteriormente. Porém existem situações nas quais podemos tolerar uma diferença entre os dados reconstruídos e os originais, sem perda significativa da qualidade. Quando toleramos esta diferença em algoritmos de compressão, podemos usar técnicas de compressão com perdas. Com estas técnicas podemos obter uma maior compressão.

2.1.2 Compressão com perdas

Nas técnicas de compressão com perdas são admitidas diferenças entre os dados recuperados e os presentes à entrada. A vantagem delas é que podemos obter uma maior taxa de compressão em relação as técnicas de compressão sem perdas.

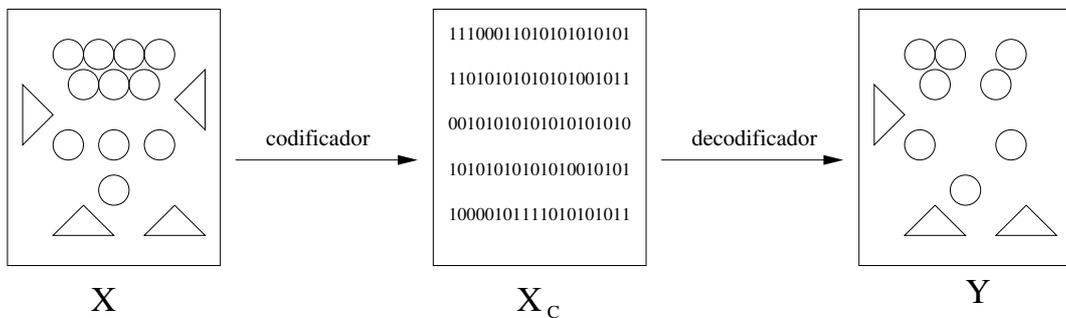


Figura 2.3: Esquema de Técnica de Compressão com perdas

A medida de desempenho das técnicas de compressão sem perdas é a taxa de compressão (ver seção 2.4). Esta medida é a principal preocupação do projetista quando está projetando um algoritmo de compressão sem perdas. Porém, para compressão com perdas, onde é permitida a diferença entre os dados presentes na entrada e os dados recuperados (diferença entre X e Y figura 2.3); a sua medida de desempenho é a distorção D . O estudo da relação entre a taxa R e a distorção D , é chamado de teoria da taxa-distorção.

Entre as técnicas de compressão com perdas, podemos citar a *quantização usando codificadores por entropia* e a *codificação por transformadas*.

A quantização usando codificadores por entropia divide a faixa de valores assumidos pela fonte de entrada, em intervalos, onde cada intervalo é representado por um código diferente, e para cada código gerado existe um valor de reconstrução. Esse método é dividido em duas etapas: *quantização* e *codificação*.

Os codificadores por transformada consistem em efetuar uma transformação da fonte de entrada em uma outra seqüência, onde a maior parte da informação está concentrada em poucos elementos, ou seja, o resultado da transformação resulta numa compactação de energia. Estes tipos de codificadores foram inicialmente utilizados por Huang e Schultheiss [7]. Este método consiste em três etapas: *transformação*, na qual se transforma uma seqüência de entrada $\{x_n\}$ em outra seqüência $\{\theta_n\}$; *quantização*, que consiste na quantização da seqüência transformada; e por último a *codificação*.

A transformação, como mencionado anteriormente, consiste em transformar uma seqüência de entrada $\{x_n\}$ em outra seqüência $\{\theta_m\}$. Essa transformação pode ser linear de acordo com a equação abaixo:

$$\theta_n = \sum_{i=0}^{N-1} x_i a_{n,i}. \quad (2.1)$$

A seqüência de entrada $\{x_n\}$ pode ser recuperada da seqüência transformada pela transformada inversa:

$$x_n = \sum_{i=0}^{N-1} \theta_i b_{n,i}. \quad (2.2)$$

E pode ser representada na forma matricial sendo:

$$\theta = \mathbf{A}\mathbf{x} \quad (2.3)$$

$$\mathbf{x} = \mathbf{B}\theta \quad (2.4)$$

Quando trabalhamos com compressão de imagens se usa em geral transformadas ortogonais ou biortogonais. Estas transformações são ditas separáveis quando podemos transformar um bloco bi-dimensional começando por uma das dimensões, e repetindo a operação ao longo da outra direção. Em termos de matrizes, transformamos primeiramente suas linhas e, após, suas colunas. E é dada na forma matricial por

$$\Theta = \mathbf{A}\mathbf{X}\mathbf{A}^T \quad (2.5)$$

E a sua transformada inversa é

$$\mathbf{X} = \mathbf{B}\Theta\mathbf{B}^T \quad (2.6)$$

Exemplos de transformadas são: transformada discreta de Walsh-Hadamard (DWHT) [8]; transformada do cosseno [9]; transformada wavelet [10, 11, 12].

A quantização consiste em representar a grande quantidade de elementos gerados pela transformação usando uma quantidade menor de símbolos. Os quantizadores são de dois tipos: *quantizador escalar* e o *quantizador vetorial*. Chamamos de quantizador escalar quando o conjunto de entrada tem valores escalares, e o vetorial quando o conjunto de entrada são vetores. Estes métodos possuem duas etapas: *quantização* e *codificação*.

A codificação consiste em codificar os símbolos gerados na quantização usando alguma técnica de codificação por **entropia**, como por exemplo, o código de Huffman [2] ou a codificação aritmética [4, 13].

Um exemplo de aplicação para compressão de imagem com perdas é o padrão JPEG [14]. A aproximação recomendada no padrão JPEG é a codificação por transformada usando a DCT.

2.2 Teoria da informação para compressão sem perdas

Sabe-se que as técnicas de compressão envolvem perdas ou não da informação, neste sentido Shannon [15] desenvolveu o primeiro estudo para verificar de forma quantitativa tal informação. Neste estudo, Shannon usou o termo *auto-informação* que pode ser definido da seguinte maneira:

Tendo-se um evento aleatório A e sua probabilidade associada $P(A)$, então podemos definir a auto-informação associada a A , por:

$$i(A) = \log_b \frac{1}{P(A)} \quad (2.7)$$

ou em bits,

$$i(A) = \log_2 \frac{1}{P(A)} \quad (2.8)$$

Pode-se concluir que se a probabilidade do evento é baixa, a quantidade de auto-informação é alta, e se a probabilidade de um evento é alta, então a auto-informação é baixa.

Shannon também contribuiu neste estudo desenvolvendo a chamada auto-informação média, ou entropia [3, 16, 17, 18] e pode ser definida da seguinte maneira:

Seja $A = \{a_1, a_2, \dots, a_n\}$, chamado alfabeto da fonte S , onde a_i são símbolos, e a saída é igual a $S = \{s_1, s_2, \dots, s_m\}$ onde todos $s_i \in A$ e suas probabilidades associadas são $P(s_i)$. A entropia, em bits/símbolo é definida por:

$$H(S) = \sum_{l=1}^m P(s_l) i(s_l). \quad (2.9)$$

Substituindo a equação 2.8 na equação 2.9

$$H(S) = - \sum_{l=1}^m P(s_l) \log_2 P(s_l) \quad (2.10)$$

2.3 Teoria da informação para compressão com perdas

Na compressão sem perdas a informação decodificada é idêntica à informação antes da compressão. Há uma preocupação com a taxa de compressão. No caso da compressão com perdas existe a preocupação com a taxa de compressão e a distorção. A distorção expressa a diferença entre a saída Y e a entrada X (figura 2.3). Ambas a taxa e a distorção podem ser representadas pela função taxa-distorção ($R(D)$).

A função taxa-distorção relaciona a distorção à taxa no esquema de compressão com perdas. Conforme visto anteriormente, a taxa é o número médio de bits para representar cada amostra. Para entendermos melhor a função taxa-distorção precisamos de dois conceitos: *entropia condicional* e *informação mútua média*.

Dadas duas variáveis aleatórias [19] X e Y , dois alfabetos $X' = \{x_1, x_2, \dots, x_n\}$ e $Y' = \{y_1, y_2, \dots, y_m\}$, e sabendo que suas entropias são

$$H(X) = - \sum_{i=0}^{N-1} P(x_i) \log_2 P(x_i) \quad (2.11)$$

e

$$H(Y) = - \sum_{j=0}^{M-1} P(y_j) \log_2 P(y_j) \quad (2.12)$$

Podemos dizer que a entropia condicional é definida segundo [17] como:

$$H(X|Y) = - \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P(x_i|y_j) P(y_j) \log_2 P(x_i|y_j) \quad (2.13)$$

Estes conceitos são necessários para podermos definir a informação mútua média em termos de entropia. Para encontrarmos a informação mútua média, conceito que necessitamos para definir a teoria taxa-distorção, precisamos lembrar do conceito de informação mútua [17] que é dada por

$$i(x_k; y_j) = \log \left[\frac{P(x_k|y_j)}{P(x_k)} \right] \quad (2.14)$$

Logo a informação mútua média é dada, em bits/símbolo, por

$$I(X; Y) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P(x_i, y_j) \log \left[\frac{P(x_i|y_j)}{P(x_i)} \right] \quad (2.15)$$

$$I(X; Y) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P(x_i|y_j)P(y_j) \log \left[\frac{P(x_i|y_j)}{P(x_i)} \right] \quad (2.16)$$

Usando as equações 2.11 e 2.12 temos

$$I(X; Y) = H(X) - H(X|Y) \quad (2.17)$$

Com base nestes conceitos podemos definir a teoria taxa-distorção. A teoria taxa-distorção conforme já mencionado, procura ver a relação satisfatória entre a taxa de compressão R e a distorção D . Esta teoria especifica a menor taxa na qual a saída da fonte pode ser codificada enquanto mantém a distorção menor ou igual a D . A função taxa-distorção é definida definida por [17] como:

Teorema 2.3.1 *A função taxa-distorção para uma fonte X identicamente distribuída com distribuição $P(x_i)$ e uma função de distribuição limitada $d(x_i, y_j)$ é o mínimo da informação mútua. Deste modo*

$$R(D) = \min_{P(x_i, y_j) | \sum_i \sum_j P(x_i)P(y_j|x_i)d(x_i, y_j) \leq D^*} I(X; Y) \quad (2.18)$$

Onde a distorção média é $D = \sum_i \sum_j P(x_i)P(y_j|x_i)d(x_i, y_j)$ e D^* é a distorção alvo.

2.4 Medidas de desempenho

Nesta seção vamos conceituar alguns critérios para medidas de desempenho de algoritmos de compressão. Os critérios serão todos voltados para esquemas de compressão com perdas para imagens de níveis de cinza codificadas com 255 níveis diferentes. Serão adotados os critérios para o esquema de compressão com perdas pelo fato do trabalho realizar os testes de simulações para este tipo de esquema de compressão.

Uma das primeiras medidas é a taxa de compressão, que é a relação entre o número de bits requeridos para representar a imagem antes da compressão (Tamanho da Imagem Original) ao número de bits necessários para representar a imagem após a compressão (Tamanho da Imagem Atual), como mostra a equação 2.19:

$$CR = \frac{\text{Tamanho da Imagem Original}}{\text{Tamanho da Imagem Atual}} \quad (2.19)$$

A taxa de compressão também pode ser representada pelo número de bits necessários para representar uma amostra simples (pixel). Isto é geralmente referido como taxa (R). Esta taxa pode ser obtida pela razão entre o tamanho da imagem depois da compressão (Tamanho Imagem Atual) e o número de pixels da imagem (Número de Pixels) (equação 2.20).

$$R = \frac{\text{Tamanho da Imagem Atual}}{\text{Número de Pixels}} \quad (2.20)$$

Como na compressão com perdas a imagem antes da compressão é diferente da imagem comprimida, torna-se necessário quantificar esta diferença. Esta diferença é frequentemente chamada de distorção, que é a outra medida que será discutida nesta seção. Neste trabalho a medida de distorção é a *peak-signal-to-noise-ratio* (PSNR), dada por:

$$PSNR = 10 \log_{10} \frac{x_{peak}^2}{\sigma_d^2}. \quad (2.21)$$

Onde:

- (x_{peak}) é o valor máximo de nível de cinza da imagem.
- (σ_d^2) é o erro médio quadrático.

Se $\{x_n\}$ é a fonte de entrada, no caso do trabalho, uma imagem, e sua reconstrução $\{y_n\}$ possui dimensões (N,M) então σ_d^2 vale (ver [3, 16]).

$$\sigma_d^2 = \frac{1}{NM} \sum_{n=1}^{NM} (x_n - y_n)^2 \quad (2.22)$$

logo a PSNR pode ser calculada como:

$$PSNR = 10 \log_{10} \frac{x_{peak}^2}{\frac{1}{NM} \sum_{n=1}^{NM} (x_n - y_n)^2}. \quad (2.23)$$

Onde $x_{peak} = 255$, normalmente.

Capítulo 3

O MMP

Este capítulo descreve um algoritmo de compressão com perdas baseado no casamento aproximado de padrões recorrentes multiescalas, que é chamado de MMP (“Multi-dimensional Multiscale Parser”), que foi recentemente desenvolvido em [1][20]. O MMP contém um dicionário que é construído e atualizado utilizando concatenações de versões contraídas ou dilatadas de blocos de imagem previamente codificados. Nas próximas seções veremos a descrição no MMP e suas versões.

3.1 Descrição do MMP

O MMP utiliza um dicionário adaptativo constituído de elementos (blocos ou palavras) com dimensões diferentes (escalas) - como mostra a figura 3.1 - para representar um bloco de entrada X . Ou seja, verifica entre os seus elementos qual é o que melhor representa o bloco de entrada X obedecendo algum critério de medida (alvo). Podemos ajustar este critério de medida de modo que possamos efetuar uma compressão sem perdas. Isto é uma outra característica do MMP, que mostra que podemos trabalhar com ele tanto para efetuar uma compressão com perdas (a técnica que será utilizada neste trabalho, ver seção 2.1.2) quanto para efetuar compressão sem perdas (ver seção 2.1.1). O dicionário do MMP é dito adaptativo porque ele é atualizado à medida que o sinal de entrada está sendo processado. E seu número de escalas é limitado e definido em função da dimensão do bloco que será processado.

Quando o MMP está sendo usado para compressão de uma imagem, ele divide a imagem em blocos de dimensão $L \times C$, no caso deste trabalho, a imagem

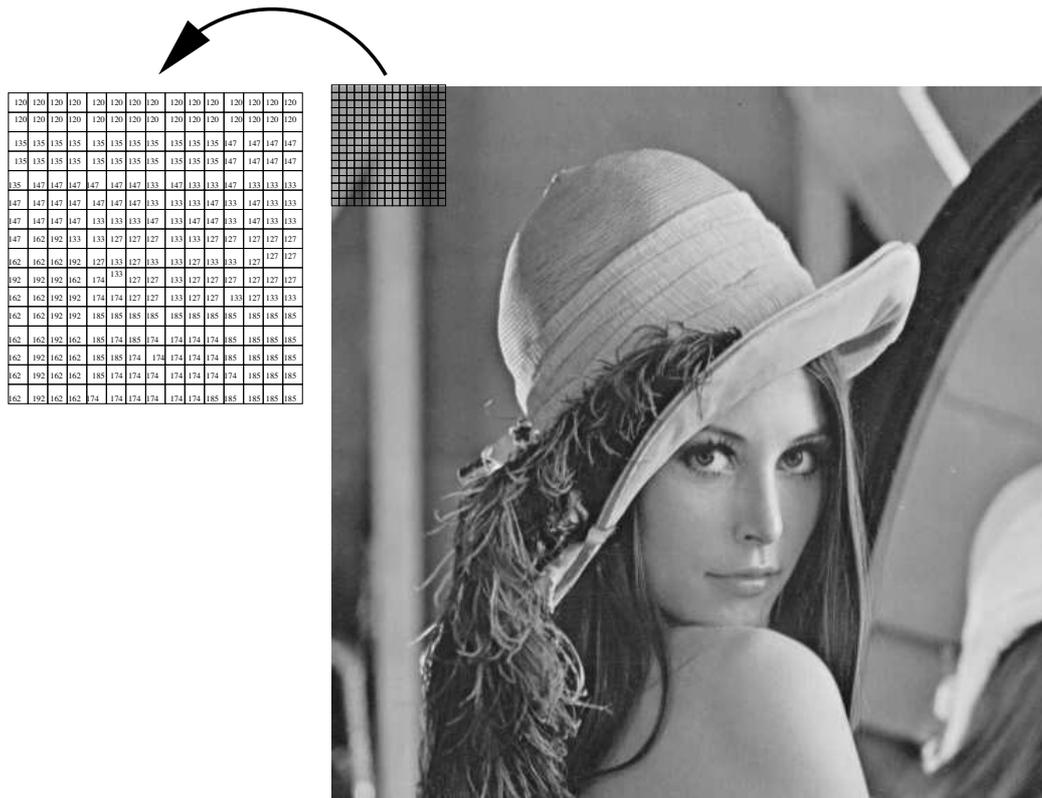
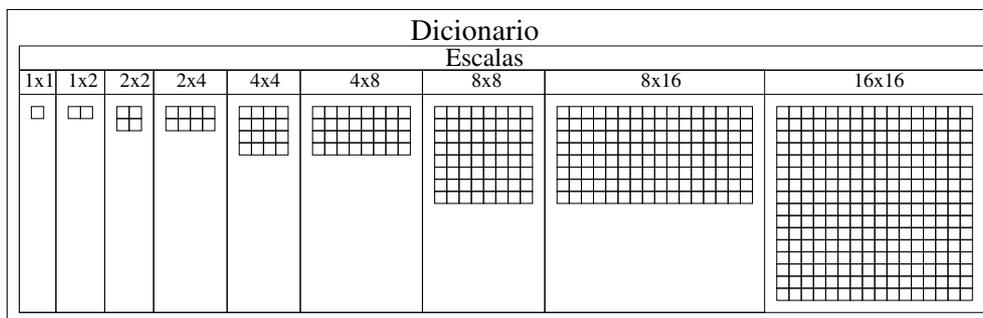


Figura 3.2: Exemplo de um bloco retirado da Imagem para processamento



206	183	127	127	127	144	144	144	168	168	183	183	206	013		
045	206	127	127	127	144	144	144	183	168	183	206	168	045	013	
045	168	183	144	144	127	127	183	144	206	168	183	206	183	045	013
206	183	144	206	183	127	183	144	168	144	183	206	183	206	013	045
045	183	206	168	144	168	127	206	168	206	183	183	206	183	013	122
045	045	206	168	168	144	168	206	183	168	168	206	183	206	248	248
248	206	183	122	013	122	045	122	045	122	013	248	248	045	122	248
206	045	206	122	013	122	013	045	122	013	045	248	248	045	144	248
045	183	183	248	248	045	122	013	045	122	122	013	144	122	013	248
183	168	183	248	013	122	045	122	013	122	013	248	248	144	122	013
183	183	168	013	248	013	122	122	045	168	013	122	248	144	144	144
206	183	119	119	119	248	168	248	248	122	122	122	144	144	144	248
119	206	003	003	119	183	122	183	122	183	122	183	183	119	119	119
119	003	119	119	003	183	119	183	183	006	183	183	003	119	119	003
003	183	183	003	119	119	119	122	183	183	248	119	119	003	119	003
003	003	119	185	206	119	248	183	003	122	122	183	183	248	144	144

Figura 3.3: Projeção do Bloco de Entrada X (parte da imagem) no dicionário

estejam disponíveis. No caso da figura 3.5, o segmento X^1 só será codificado e o dicionário atualizado quando as aproximações \hat{X}^3 e \hat{X}^4 estiverem disponíveis. A aproximação \hat{X}^1 será a concatenação de \hat{X}^3 e \hat{X}^4 . O dicionário atualizará todas as

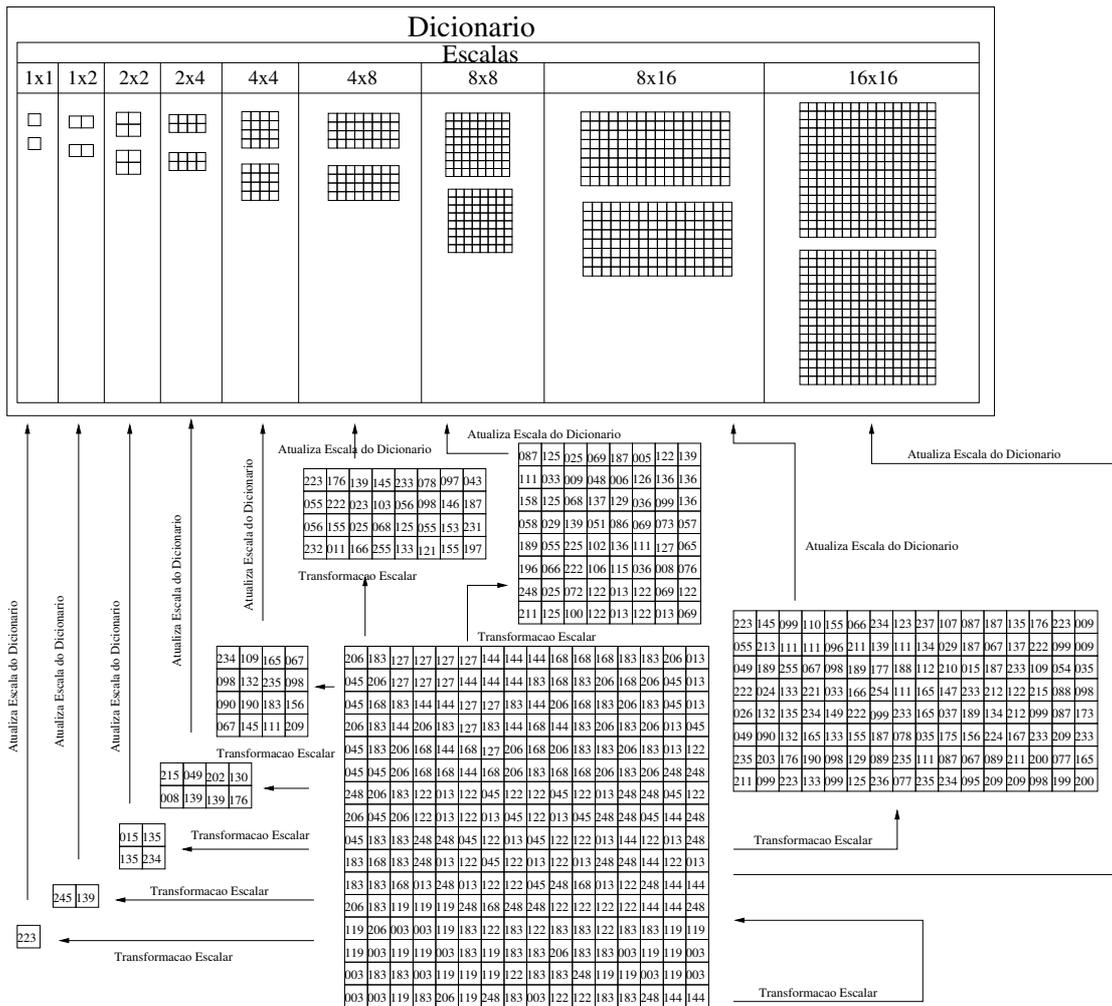


Figura 3.4: Atualização do Dicionário

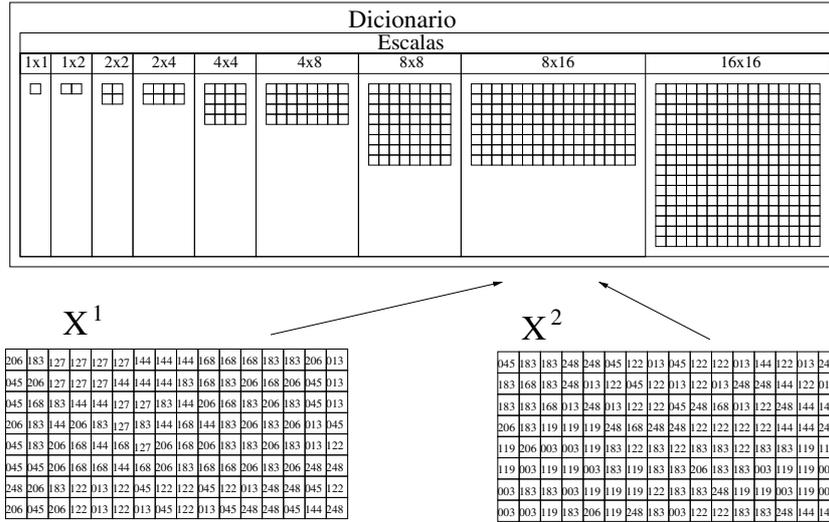


Figura 3.7: Projção dos segmentos dos blocos de entrada X (X^1 e X^2 no dicionário)

alguma aproximação no dicionário que satisfaça ao critério de medida (figura 3.3). Caso nenhum bloco do dicionário satisfaça ao critério de medida, o bloco X será segmentado em X^1 e X^2 como mostra a figura 3.5. Cada um desses segmentos será novamente escalado no dicionário e verificado se satisfaz ao critério de medida. Caso o segmento X^1 não satisfaça ao critério de medida, o mesmo será dividido em X^3 e X^4 (figura 3.6). Os segmentos X^3 e X^4 passam pelo mesmo processo que o segmento X^1 . Se os segmentos X^3 e X^4 satisfazem ao critério de medida, obtem-se as aproximações \hat{X}^3 e \hat{X}^4 . Uma vez satisfeito o critério os segmentos \hat{X}^3 e \hat{X}^4 serão codificados cada um e a aproximação de X^1 será a concatenação de \hat{X}^3 e \hat{X}^4 . Após a codificação das aproximações \hat{X}^3 e \hat{X}^4 , é a vez do segmento X^2 passar pelo mesmo processo. Como aproximação do segmento X^2 satisfazendo ao critério de medida, obtemos a aproximação \hat{X}^2 . E a aproximação \hat{X}^2 é codificada. Portanto um bloco X terá sua aproximação como mostra a figura 3.8, e através da concatenação desses segmentos obtem-se $\hat{X} = (\hat{X}^3 \hat{X}^4 \hat{X}^2)$.

O MMP codifica os blocos somente quando as aproximações atendem ao critério de medida. Ele produz um arquivo contendo seqüência de bits que representam os índices dos elementos do dicionário para cada segmento extraído do bloco de entrada, e os flags usados para informar sobre a segmentação da árvore (informa se a árvore foi segmentada ou não, ou seja, se o flag é “1” foi segmentada, se é “0” não foi segmentada). Para representar o índice do dicionário que corresponde ao bloco aproximado será sempre utilizado um código. Assim, os blocos mais utilizados e os

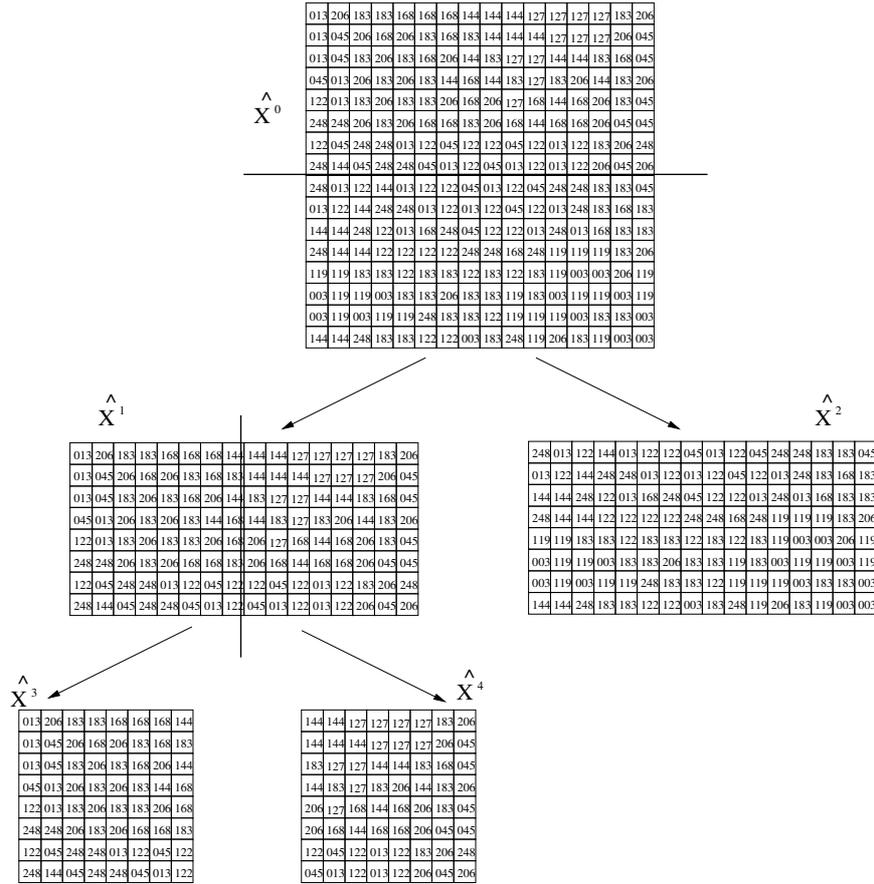


Figura 3.8: Árvore de segmentação aproximada do Bloco de Entrada X

menos utilizados serão sempre representados pelo mesmo número de bits. Portanto, teremos sempre um código de comprimento fixo. Nas próximas seções explicaremos como resolvemos este problema.

3.2 Versões do MMP

Neste trabalho estuda-se o comportamento de várias versões do MMP, do básico (já explicado nas seções anteriores) até a otimização taxa-distorção. Em cada uma das versões, o MMP passou por modificações que melhoraram seu desempenho na compressão de imagens [1][21]. Cada fase deste processo é comentada a seguir, até chegarmos a versão proposta no trabalho.

3.2.1 MMP com Codificador Aritmético

Com o propósito de eliminar as redundâncias na codificação, melhorar a taxa de bits e ter o código de comprimento variável, o MMP foi modificado utilizando a codificação aritmética [13]. Na codificação aritmética a informação é representada por um intervalo de números reais entre 0 e 1. Como a informação é longa, o intervalo necessário para representá-la é pequeno, e o número de bits necessários para especificar este intervalo cresce. Sucessivos símbolos da informação reduzem o tamanho do intervalo de acordo com as probabilidades dos símbolos geradas pelo modelo, que pode ser fixo ou adaptável. Nesta fase os flags, que indicam se há segmentação ou não do bloco, e os índices do dicionário, são codificados através da codificação aritmética. Então é verificada a frequência (probabilidade) em que ocorre um determinado índice do dicionário e representá-lo com o número de bits necessários.

3.2.2 MMP Otimização Taxa-distorção (MMPRD)

Nesta fase do MMP houve uma otimização na árvore de segmentação. Esta otimização é chamada de Otimização Taxa-Distorção [1], que otimiza a árvore de segmentação preocupando-se tanto com a taxa de bits quanto com a distorção. Nas fases anteriores, o MMP preocupava-se somente com a distorção alvo D , para decidir se um bloco será codificado ou não. Neste caso, é analisado cada nodo da árvore de segmentação obedecendo um critério de fidelidade chamado “custo” ($J = D + \lambda R$). Esta análise depende dos nodos pais e filhos. A próxima subseção explicará com mais detalhes esta otimização.

3.2.3 Otimização taxa-distorção

Como mencionado nas seções anteriores, tínhamos o MMP controlado por um critério de distorção D . A segmentação do bloco de entrada X era interrompida quando a distorção alvo no bloco aproximado (\hat{X}) do bloco X era menor ou igual a D . Com o propósito de melhorar o desempenho de decisões da segmentação do bloco, o critério passou a ser baseado na otimização da segmentação da árvore segundo um critério baseado tanto na taxa quanto na distorção.

3.2.3.1 Otimização da árvore de segmentação

O MMP é um algoritmo de compressão de dados com perdas que utiliza um dicionário em S e segmentação de um bloco de entrada X . O bloco é dividido em L segmentos de acordo com o elemento do dicionário S também de comprimento L . Cada segmento do bloco é dividido e codificado conforme um elemento do dicionário S escolhido para o tamanho apropriado, que atenda ao critério de medida, como descrito neste capítulo. Assim obtemos o bloco aproximação de X (\hat{X}). A segmentação resultante pode ser representada por uma árvore de segmentação, conforme figura 3.8, ou também podemos representá-la como na figura 3.9.

Cada segmento utilizado para representar um bloco aproximado de X (\hat{X}), pode ser considerado como nodo folha da árvore sem descendentes. Podemos observar pela figura 3.9 que o bloco de entrada X foi segmentado em 3 segmentos, ou seja, $X = (X_3 X_4 X_2)$. Cada um dos segmentos foi aproximado por um bloco do dicionário S , com alguma distorção, e podemos associar cada nodo folha da árvore a um valor de distorção, e cada nodo está associado algum bloco de X . Por exemplo, o nodo n_1 está associado a um segmento X^1 que não foi utilizado para representar o segmento X^1 , porque este foi segmentado em $X^1 = (X^3 X^4)$. Portanto podemos associar a cada nodo um a distorção referente a este nodo $D(n_l)$. E também definimos $R(n_l)$ como a taxa necessária para especificar o índice i_l do dicionário S , ou seja,

$$R(n_l) = -\log_2(Pr(i_l)) \quad (3.1)$$

O critério, neste caso, será agora calculado para cada nodo através do custo lagrangeano [22, 20, 1], que define a relação entre a distorção e a taxa para um nó em particular:

$$J(n_l) = D(n_l) + \lambda R(n_l). \quad (3.2)$$

Onde:

- $J(n_l)$ é o custo do nodo n_l .
- $D(n_l)$ é a distorção entre o bloco de entrada X associado ao nó n_l e a sua aproximação do dicionário S com a mesma dimensão do bloco de entrada X .

- $R(n_l)$ é a taxa para representar o índice i_l do dicionário S .
- λ é o fator ponderador entre a taxa e distorção.

A otimização pelo método de Lagrange foi primeiramente introduzida em [23], podemos encontrar aplicações deste método em [24, 25, 26]

Agora ao invés de verificarmos a distorção num determinado nodo, o que será verificado é o custo deste nodo (n_l).

Um multiplicador de Lagrange é escolhido e a partir deste valor calcula-se o custo do nodo. Vale lembrar que neste momento calculamos tanto o custo do nodo folha, pai, (n_l) quanto dos seus nodos descendentes, filhos, (n_{2l+1} e n_{2l+2}). Verifica-se o custo do nodo pai (n_l), através do multiplicador de Lagrange, distorção e taxa do índice, e dos nodos filhos. Lembrando que são verificados os blocos aproximações do dicionário atendendo agora a este novo critério de fidelidade. Caso o custo do nodo pai (n_l), $J(n_l) = D(n_l) + \lambda R(n_l)$, seja menor ou igual ao custo dos nodos filhos + a taxa do flag de segmentação, como mostra a equação 3.3, o nodo pai mantém-se e os nodos filhos da árvore são descartados, assim o nodo pai passa a ser um nodo folha ou um dos nodos filhos de um nodo folha ascendente. Caso contrário, os nodos filhos mantêm-se (figura 3.9).

$$J(n_l) < J(n_{2l+1}) + J(n_{2l+2}) + \lambda R(flag(n_l)) \quad (3.3)$$

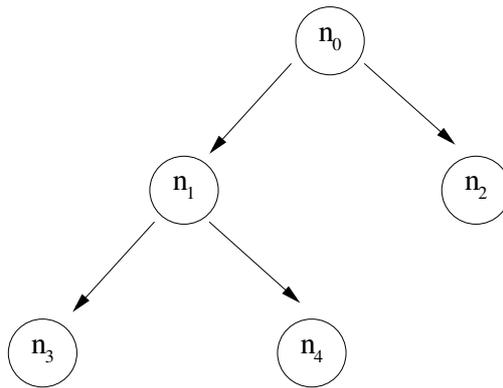


Figura 3.9: Árvore de segmentação representada com nodos folhas

3.2.4 MMP com Otimização taxa-distorção intermediário (MMPRDI)

Esta é a última fase implementada do MMP. Na fase anterior não nos preocupávamos com as frequências dos flags e dos índices quando era calculado os custos dos nodos. Agora há uma preocupação com estas frequências, e por isso, um dicionário rascunho foi criado para ajudar a verificação destas frequências para melhorar a codificação da imagem [21]. Neste algoritmo o dicionário D é simulado através do dicionário rascunho e todas as atualizações serão simuladas e armazenadas nele. O MMPRDI inicia uma árvore com todos os possíveis nós folhas. A árvore é moldada seguindo-se um critério. O critério é a minimização do custo lagrangeano conforme seção 3.2.3.1 (equação 3.3). Os custos são verificados até obtermos um melhor molde da árvore que minimize o custo lagrangeano total. Determinado o melhor molde da árvore efetuamos a codificação. O trabalho desenvolvido nesta tese inicia-se a partir desta versão do MMP. Faremos um processo de análise do desempenho de codificação de esquemas de compressão de sinais baseados no MMP, com modificações que possam redundar em uma melhora de desempenho. As modificações estão relacionadas à construção e atualização do dicionário. Elas consistem em segmentar o dicionário no espaço de códigos (capítulo 4) e, além disso, incluir no dicionário blocos rotacionados e transpostos (capítulo 5), e blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho (capítulo 6).

Capítulo 4

MMPRDI usando segmentação do dicionário do espaço de códigos

O MMPRDI [21] utiliza dicionários constituídos de blocos de tamanhos diferentes (um dicionário para cada escala) que são atualizados com expansões e contrações de blocos concatenados através da transformação de escala do bloco aproximado (ver seção 4.2). Este capítulo descreve o MMPRDI com uma modificação na estrutura dos dicionários e na codificação de seus índices. Os dicionários que antes eram como na figura 3.1 serão segmentados, ou seja, os dicionários em cada escala serão divididos em subdicionários como mostra a figura 4.1. Espera-se que com esta segmentação o MMPRDI possa ter um melhor desempenho na codificação de imagens. No MMPRDI antes desta modificação, os dicionários em cada escala eram todos atualizados com expansões e contrações de segmentos previamente codificados, não importando de qual escala vinha o segmento para a atualização do dicionário. Ou seja, não havia a preocupação com a origem do bloco usado para atualizar o dicionário. Agora, o dicionário de cada escala é dividido em subdicionários. É o subdicionário que informa, para cada elemento incluído no dicionário, à qual escala ele pertence antes da transformação de escala. Por exemplo, suponhamos que a palavra do dicionário que mais aproxima-se do bloco de entrada X tem tamanho 2×2 (escala do dicionário). Esta palavra passa pela transformação de escalas, resultando em expansões e contrações para atualizar o dicionário. Na hora da atualização de cada escala do dicionário, estas expansões e contrações são colocadas justamente no subdicionário 2×2 de cada escala do dicionário. Então, podemos saber qual

escala do dicionário que mais contribuiu na atualização do mesmo. Quanto à codificação, a palavra do dicionário era codificada diretamente com um flag informando sobre a segmentação da árvore e um outro informando o índice do dicionário. Agora, na segmentação de espaço de códigos, ao invés de codificar o índice do dicionário, codifica-se um flag informando a qual subdicionário pertence a palavra, e depois um índice indicando a sua localização dentro do subdicionário. A codificação utiliza um codificador aritmético com um modelo para cada escala do dicionário. A figura 4.2 mostra como será agora a estrutura do dicionário. As próximas seções falam da motivação, apresentam a descrição deste algoritmo, bem como os resultados do levantamento estatístico das palavras do dicionário e as curvas taxa-distorção para segmentação do dicionário no espaço de códigos.

Dicionário

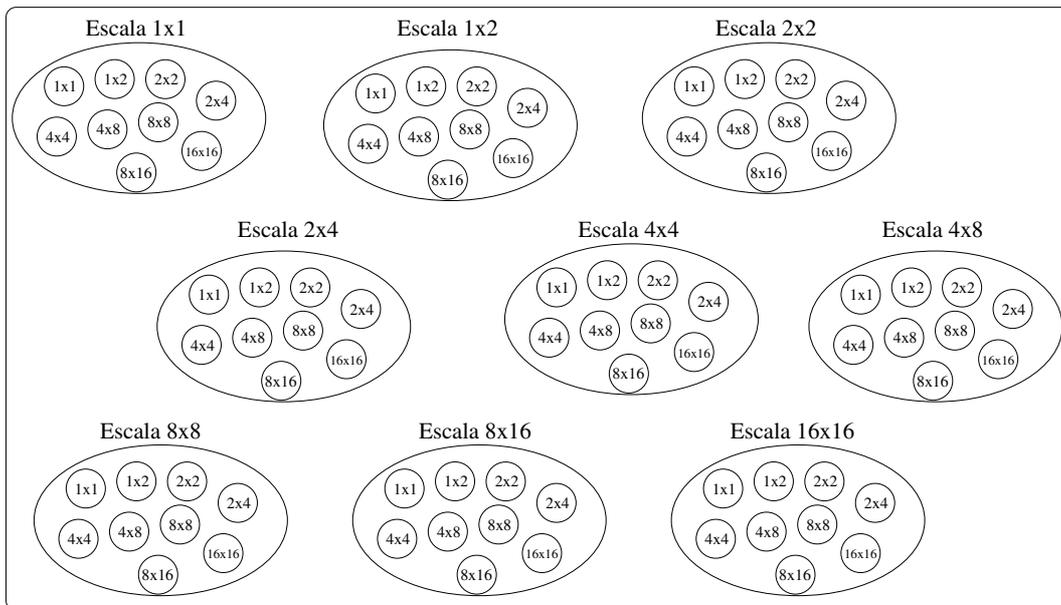


Figura 4.1: Novo esquema do dicionário para avaliação do desempenho do MMRDI

4.1 Motivação

A motivação principal neste trabalho é verificar se é efetivo ou não o uso da segmentação no espaço de códigos. Para tal será feito um levantamento estatístico das probabilidades dos vetores de cada subdicionário. Note que será vantajoso o uso da segmentação quando as probabilidades destes vetores forem bastante diferentes

<i>Dicionario</i>									
Subdicionario	Escala								
	1x1	1x2	2x2	2x4	4x4	4x8	8x8	8x16	16x16
1x1									
1x2									
2x2									
2x4									
4x4									
4x8									
8x8									
8x16									
16x16									

Figura 4.2: Estrutura do dicionário para avaliação do desempenho do MMRDI

entre si (ver seção 4.4). A partir daí, será avaliado o desempenho taxa-distorção do MMRDI usando a segmentação no espaço de códigos.

4.2 Descrição

O algoritmo que será usado neste capítulo é o MMRDI (ver capítulo 3) com uma modificação na estrutura do dicionário e na codificação dos seus índices. O bloco de entrada X é dividido em L segmentos resultando em uma árvore de segmentação na qual cada nó folha corresponde a um elemento do dicionário. A árvore é moldada atendendo a um critério, que neste caso é a minimização do custo

lagrangeano (ver capítulo 3). Com este critério verifica-se se o custo lagrangeano do nó pai é menor ou igual a somas dos custos dos nós filhos. Caso o custo do nó pai seja menor ou igual a soma dos custos dos nós filhos, os nós filhos são descartados, o nó pai mantém-se e passa a ser um nó folha, ou seja, um dos nós filhos de seu nó ascendente. Os custos são verificados até obtermos uma árvore de segmentação ótima, ou seja, que minimize o custo lagrangeano total. Os blocos concatenados resultantes desta segmentação serão utilizados para a atualização do dicionário. A seção seguinte descreve a atualização do dicionário.

4.3 Atualização do dicionário

Dada a árvore que minimiza o custo lagrangeano, a cada nó pai corresponderá uma nova palavra nos dicionários, que consistirá da concatenação dos dois nós filhos, escaladas para o tamanho correspondente. A mesma passará por uma transformação de escala para atualizar cada escala do dicionário. O resultado de cada transformação será colocado no subdicionário correspondente ao tamanho original do nó pai em cada escala do dicionário. Por exemplo, suponhamos que este bloco com dimensão 16×16 seja transformado numa escala 8×8 do dicionário, este novo bloco será colocado na escala 8×8 no subdicionário 16×16 , pois informa que este bloco é uma transformação de escala de um bloco 16×16 que foi inserido na escala 8×8 do dicionário (figura 4.4), e assim sucessivamente. Isto permitirá que façamos a segmentação do dicionário no espaço de códigos, da seguinte forma: encontrada a palavra no dicionário que minimiza o custo lagrangeano para aquele bloco, ela será codificada com um flag informando a qual subdicionário ela pertence e um índice indicando qual a sua localização dentro do subdicionário. A codificação do flag é feita usando um codificador aritmético com um modelo para cada escala do dicionário. A codificação dos índices é equivalente à do MMPREDI. Designaremos o algoritmo por MMPREDI com segmentação no espaço de códigos (MMPREDI-seg).

4.4 Resultados

Vai valer a pena fazer a segmentação no espaço de códigos, se, em geral, em uma dada escala, as probabilidades de vetores pertencerem a um subdicionário X

<i>Dicionário</i>									
Subdicionário	Escala								
	1x1	1x2	2x2	2x4	4x4	4x8	8x8	8x16	16x16
1x1									
1x2									
2x2									
2x4									
4x4									
4x8									
8x8									
8x16									
16x16									

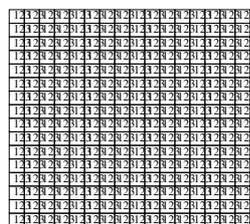


Figura 4.3: Projção do bloco de entrada X no dicionário

tenderem a ser bastante diferentes entre si. Isto acontece devido a dois fatores: (1) podemos usar um codificador aritmético adaptativo para codificar o flag que escolhe

<i>Dicionario</i>									
Subdicionario	Escala								
	1x1	1x2	2x2	2x4	4x4	4x8	8x8	8x16	16x16
1x1									
1x2									
2x2									
2x4									
4x4									
4x8									
8x8									
8x16									
16x16									

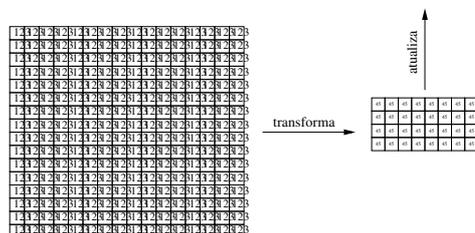


Figura 4.4: Exemplo da atualização do dicionário

entre os subdicionários. Se as probabilidades de vetores pertencerem a diferentes subdicionários forem bastante diferentes entre si, então este flag será codificado com menos de 1 bit, o que tornará a codificação eficiente. Note que se as probabilidades

dos subdicionários forem parecidas, não haverá vantagem nisto; (2) Cada subdicionário terá menos elementos, o que torna a adaptação do modelo estatístico dos seus elementos mais efetiva. Para verificar a validade deste fato, precisamos fazer um levantamento estatístico do comportamento do MMPRDI.

Aqui utilizamos o algoritmo MMPRDI descrito neste capítulo 3 para comprimir uma série de imagens de teste (apêndice A). Elas são imagens mistas e suaves. As imagens suaves são aquelas que não contém textos, no caso as imagens: aerial (apêndice A.1), baboon (apêndice A.2), barbara (apêndice A.3), bridge (apêndice A.4), gold (apêndice A.6), f16 (apêndice A.5) e lena (apêndice A.7). As imagens mistas são as que contém textos, e alguns casos, textos e figuras, como a pp1205 (apêndice A.8) e pp1209 (apêndice A.9), respectivamente. Faremos um levantamento estatístico das probabilidades das frequências dos vetores dos subdicionários do MMPRDI para cada imagem, a fim de verificar se elas são bastante diferentes entre si. Assim, poderemos analisar se é efetivo o uso da segmentação no espaço de códigos. As figuras 4.5 a 4.13 mostram os resultados estatísticos da simulação do MMPRDI com as imagens de teste. O eixo horizontal (eixo x) mostra os subdicionários 1×1 , 1×2 , 2×2 , ..., 16×16 , são eles que informam de qual escala do dicionário veio a transformação de escalas (contração ou expansão) da palavra que consiste da concatenação dos dois nós filhos da árvore de segmentação ótima (ver seção 4.2); o eixo vertical (eixo y) mostra as probabilidades dos vetores do subdicionário; e as barras verticais são as escalas do dicionário 1×1 , 1×2 , 2×2 , ..., 16×16 (como podemos ver também pela legenda). Os gráficos mostram que cada escala do dicionário possui uma porcentagem em cada subdicionário resultando em 100%. Como podemos notar pelas figuras 4.5 a 4.12 as escalas mais utilizadas para gerarem as escalas de outros dicionários, foram as escalas 4×8 . A figura 4.13 foi a única com resultado diferente que mostra que a escala do dicionário mais usada foi a 4×4 . Com base nos resultados podemos observar que as probabilidades dos vetores pertencentes aos subdicionários, no levantamento estatístico, são bastante diferentes entre si, tendo destaque o subdicionário 4×8 . Assim, podemos concluir que é útil o uso da segmentação no espaço de códigos. A próxima seção mostra o resultado da avaliação do desempenho taxa-distorção do MMPRDI com o uso da segmentação no espaço de códigos (MMPRDI-seg).

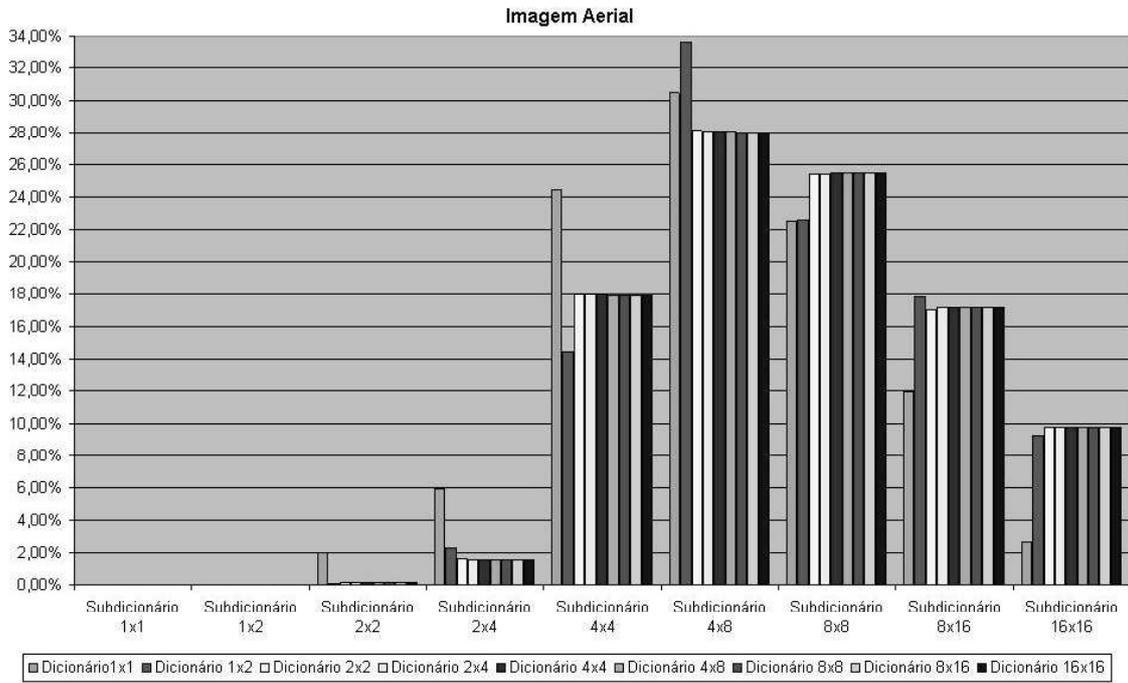


Figura 4.5: Resultado do desempenho da imagem de teste aerial (apêndice A.1).

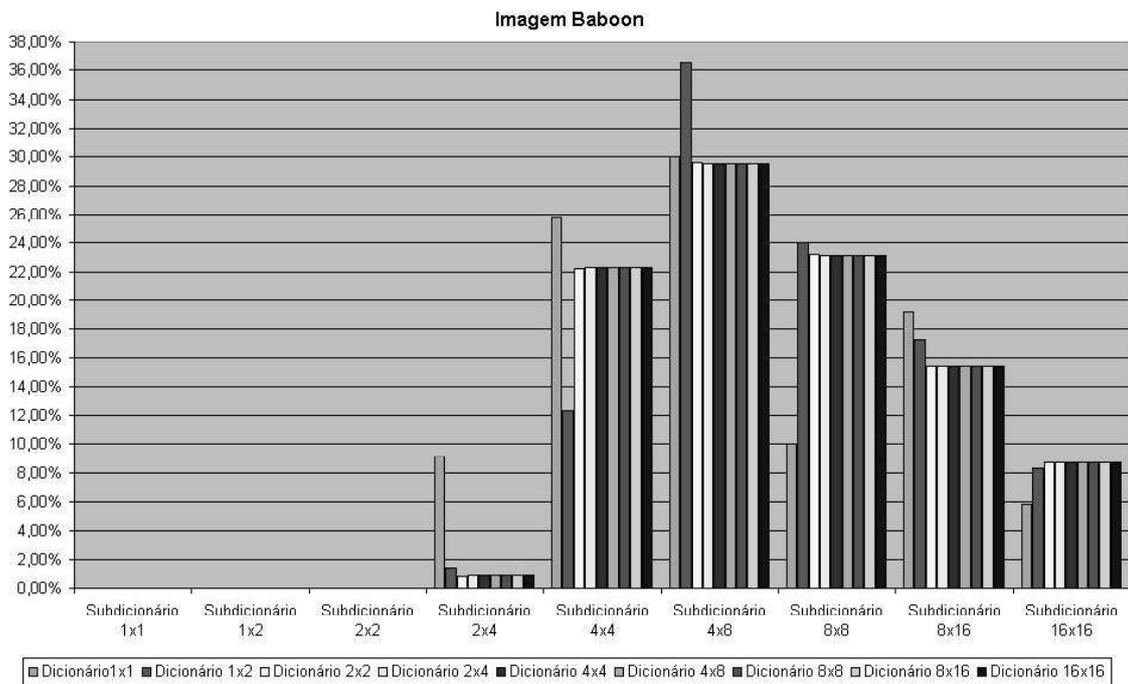


Figura 4.6: Resultado do desempenho da imagem de teste baboon (apêndice A.2).

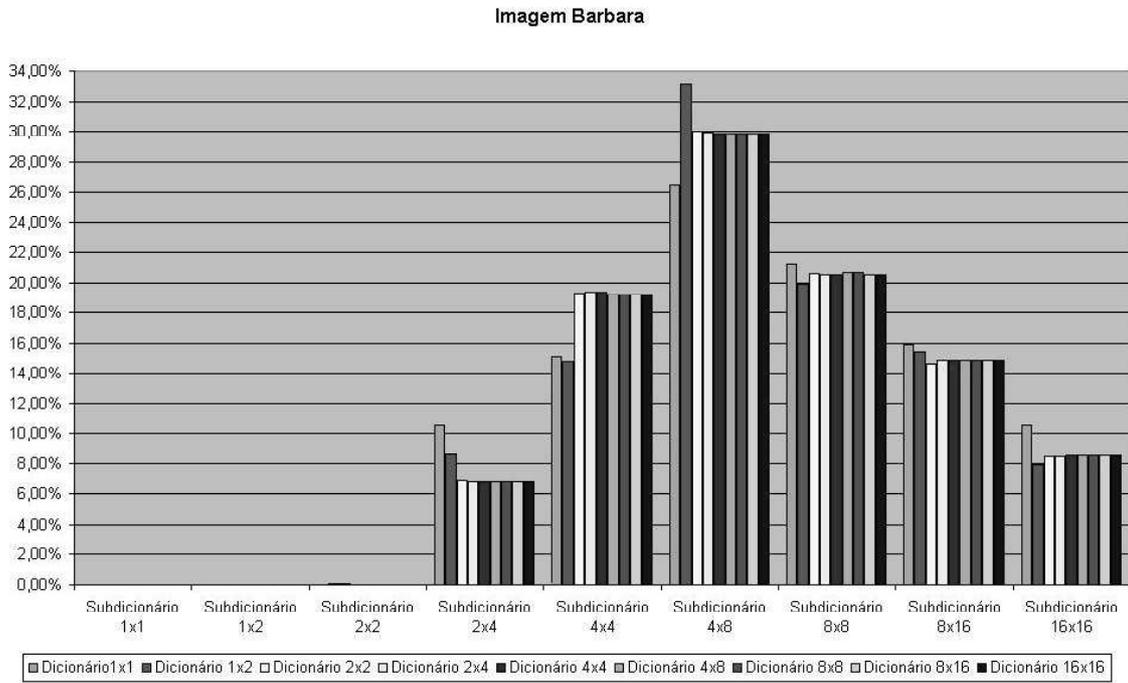


Figura 4.7: Resultado do desempenho da imagem de teste barbara (apêndice A.3).

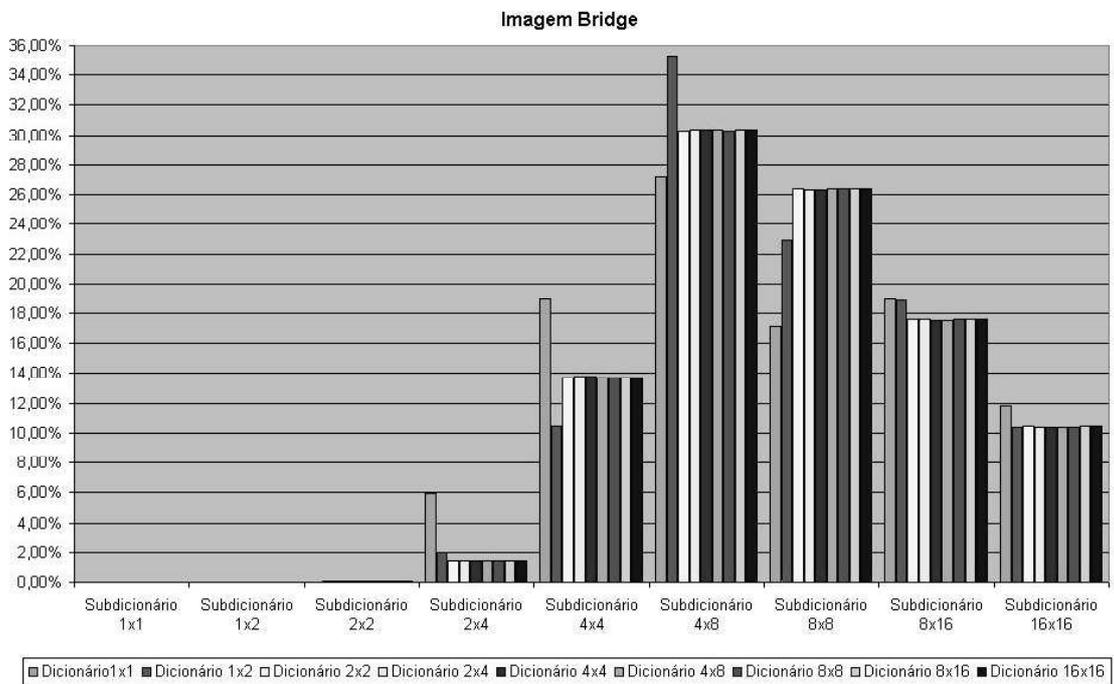


Figura 4.8: Resultado do desempenho da imagem de teste bridge (apêndice A.4).

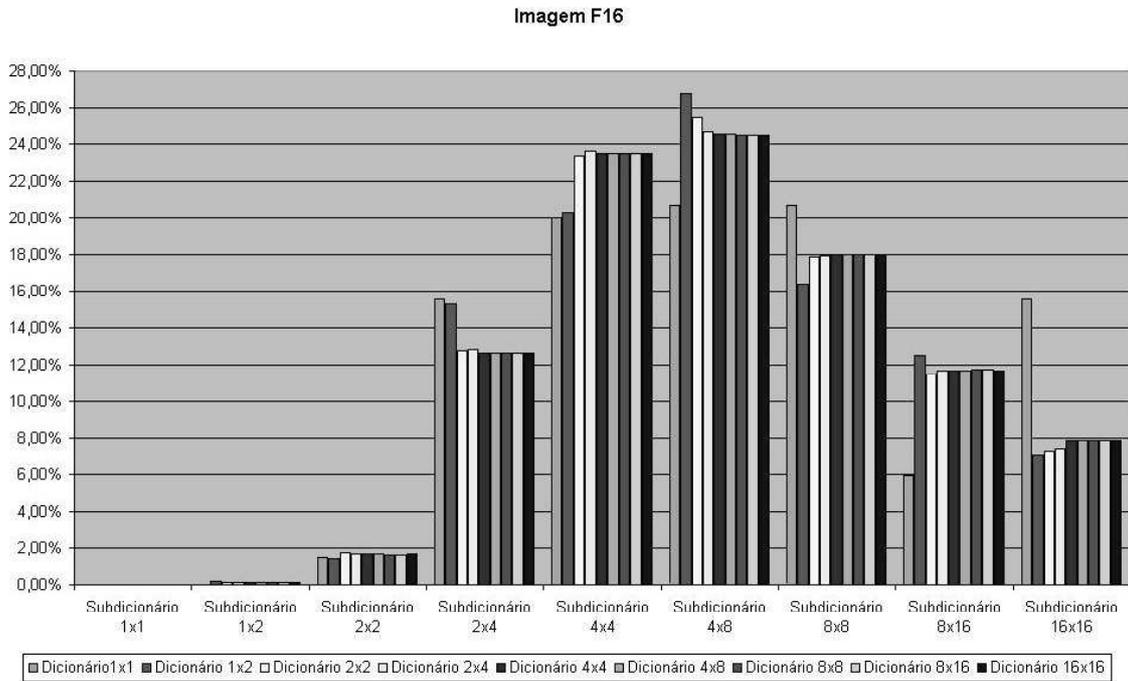


Figura 4.9: Resultado do desempenho da imagem de teste f16 (apêndice A.5).

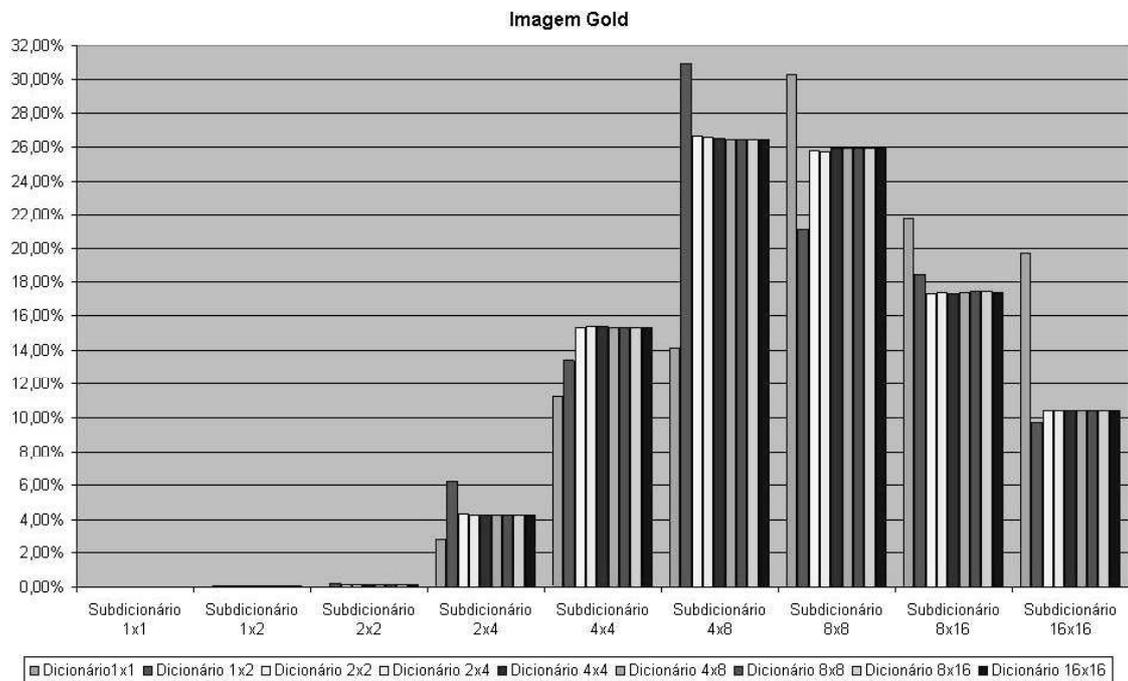


Figura 4.10: Resultado do desempenho da imagem de teste gold (apêndice A.6).

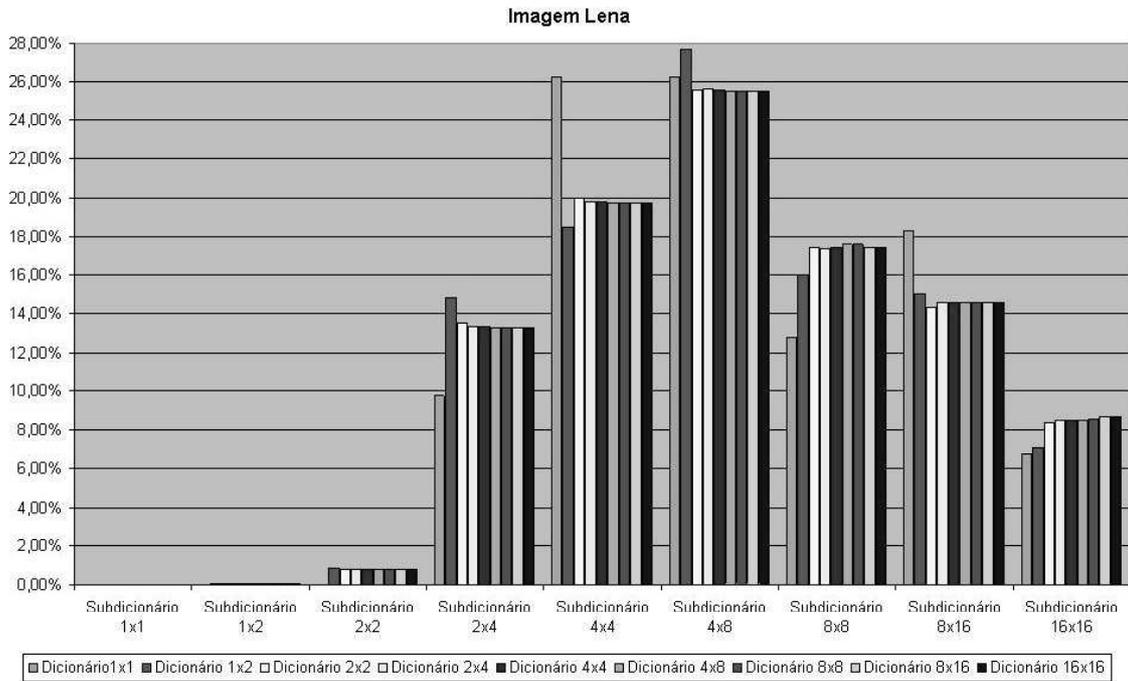


Figura 4.11: Resultado do desempenho da imagem de teste lena (apêndice A.7).

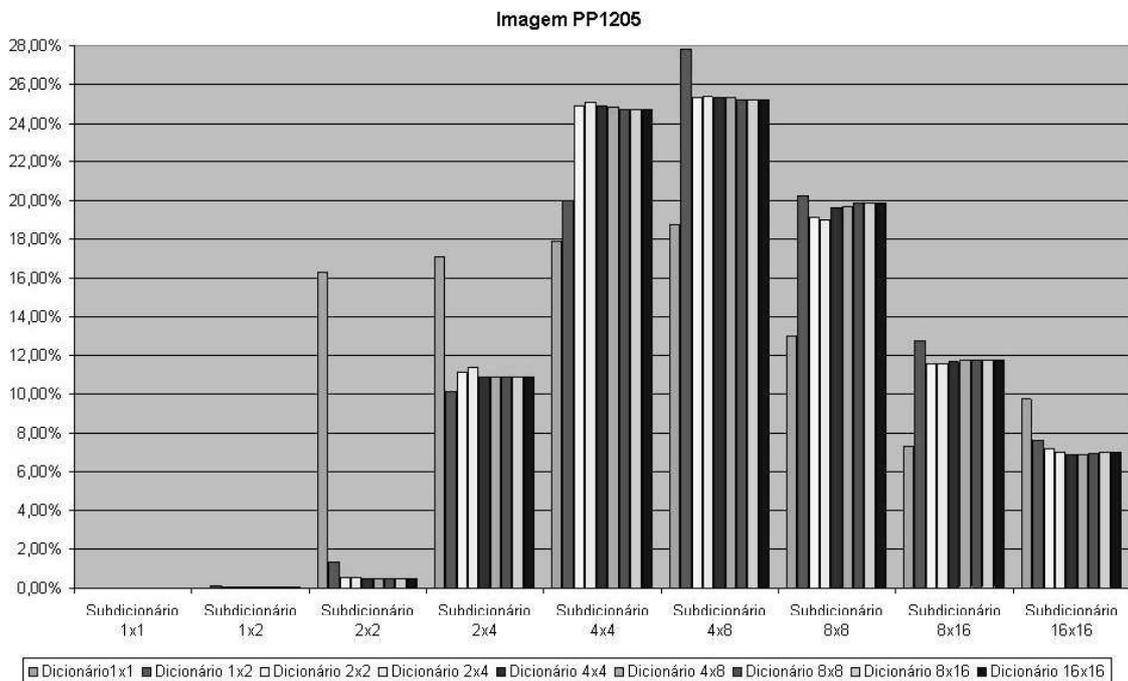


Figura 4.12: Resultado do desempenho da imagem de teste pp1205 (apêndice A.8).

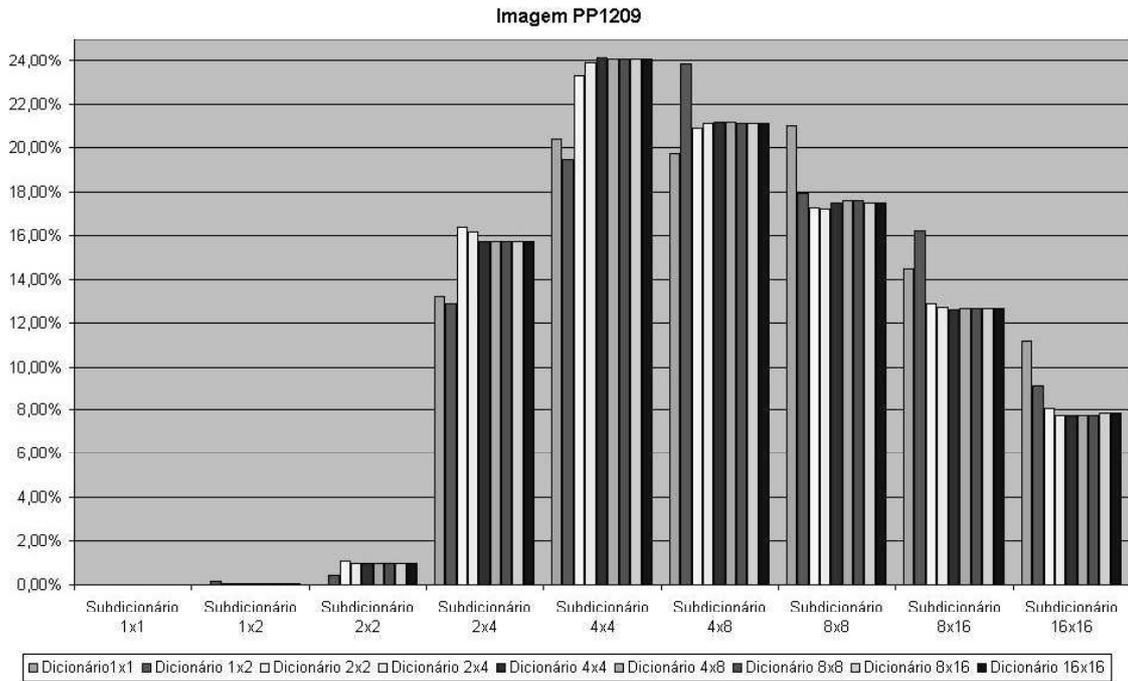


Figura 4.13: Resultado do desempenho da imagem de teste pp1209 (apêndice A.9).

4.4.1 Curvas taxa-distorção

O algoritmo MMP foi alterado para podermos verificar o desempenho taxa-distorção do MMRDI usando segmentação no espaço de códigos (MMRDI-seg). Conforme mencionado nas seções 4.2 e 4.3 o MMRDI teve a estrutura do seu dicionário alterada (figura 4.2). O dicionário agora é dividido em escalas e suas escalas divididas em subdicionários. Portanto, quando atualiza-se o dicionário, o resultado do processamento da transformação de escala será usado para atualizar cada escala do dicionário num determinado subdicionário correspondente ao tamanho da palavra antes de sua transformação de escala. Esta mesma palavra será utilizada para codificar parte da imagem. Na codificação será informado a qual subdicionário esta palavra pertence e sua localização (índice) dentro do subdicionário. Ou seja, será utilizado um flag para informar a qual subdicionário a mesma pertence e um índice indicando sua posição dentro deste subdicionário. Para codificar esta informação foi utilizado um codificador aritmético com um modelo para cada escala do dicionário.

Nesta seção mostramos os resultados em termos de curvas taxa-distorção

de simulações do MMPREDI e do MMPREDI-seg (MMPREDI usando segmentação do dicionário do espaço de códigos) usando as imagens de teste (apêndice A) como informação de entrada. As figuras 4.14 a 4.19 incluem a comparação dos resultados do MMPREDI e do MMPREDI-seg com o desempenho de algoritmos codificadores de imagens como o SPIHT [27, 28], cuja transformada é a DWT (transformada wavelet) [10, 11, 12] e JPEG [14], que usa a transformada discreta do cosseno (DCT) [9, 29]. Podemos verificar pelas figuras 4.14 a 4.17 que o desempenho do MMPREDI-seg foi melhor em relação ao MMPREDI, em boa parte das imagens houve uma melhora de aproximadamente 0.5dB a 0.5bpp, como podemos ver pela curva taxa-distorção da imagem lena (figura 4.17), e para a curva taxa-distorção da imagem barbara (figura 4.14) foi de aproximadamente 1 dB.

Já as figuras 4.18 e 4.19 mostram uma melhora em relação as outras imagens. O que podemos perceber é MMPREDI-seg para imagens mistas (imagens contendo texto e figuras) consegue uma compactação melhor que o SPIHT e JPEG. Em ambas as figuras o MMPREDI-seg teve um desempenho bem melhor que o MMPREDI. Podemos concluir que MMPREDI com segmentação do dicionário no espaço de códigos, através da curva taxa-distorção, mostra uma melhora significativa de desempenho. Nos próximos capítulos veremos as curvas taxa-distorção para o MMPREDI com a inclusão no dicionário de blocos rotacionados e transpostos, além dos blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho, para verificarmos se pode ser efetivo usarmos o MMPREDI com segmentação do dicionário no espaço de códigos.

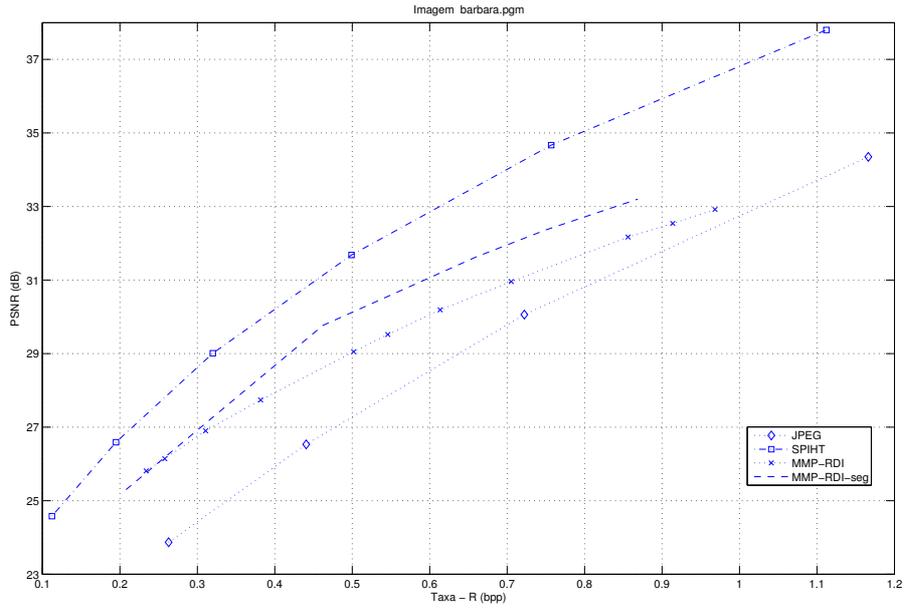


Figura 4.14: Curva Taxa-distorção do MMPRDI, MMPRDI-seg para imagem Barbara (apêndice A.3) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

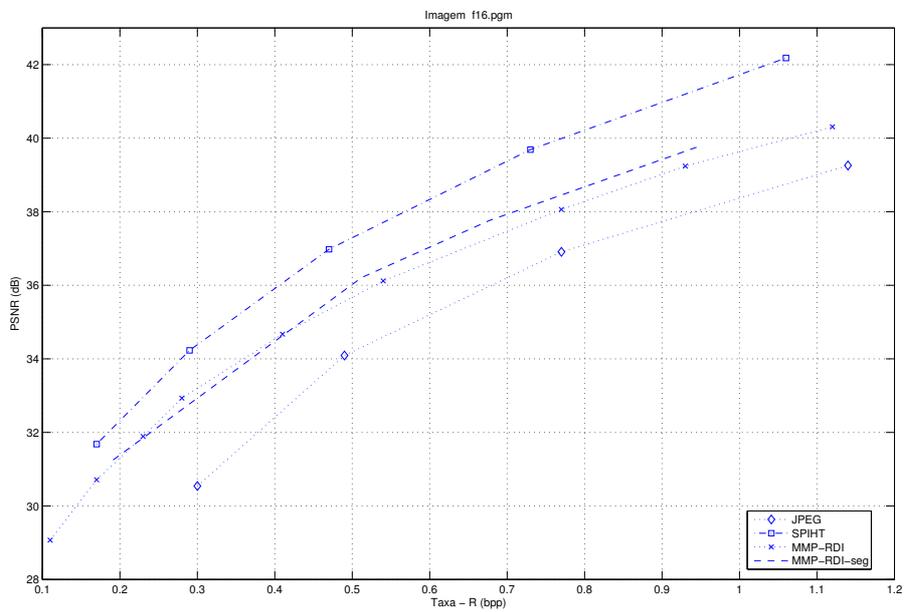


Figura 4.15: Curva Taxa-distorção do MMPRDI, MMPRDI-seg para imagem F16 (apêndice A.5) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

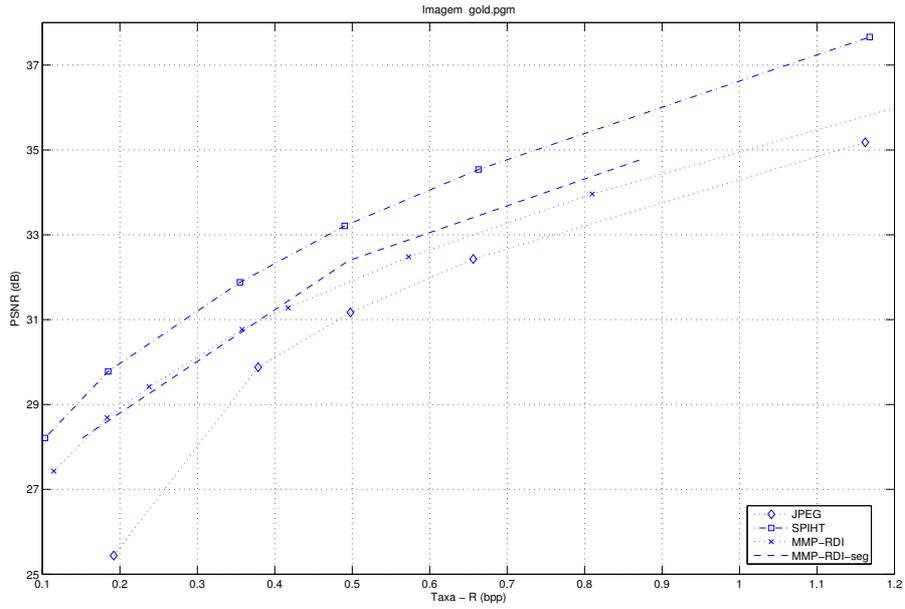


Figura 4.16: Curva Taxa-distorção do MMPRDI, MMPRDI-seg para imagem Gold (apêndice A.6) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

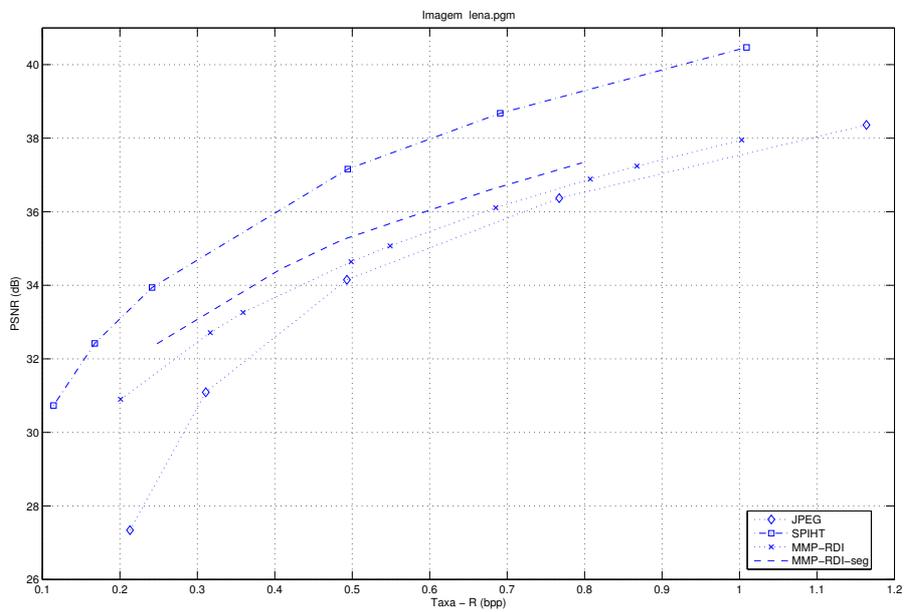


Figura 4.17: Curva Taxa-distorção do MMPRDI, MMPRDI-seg para imagem Lena (apêndice A.7) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

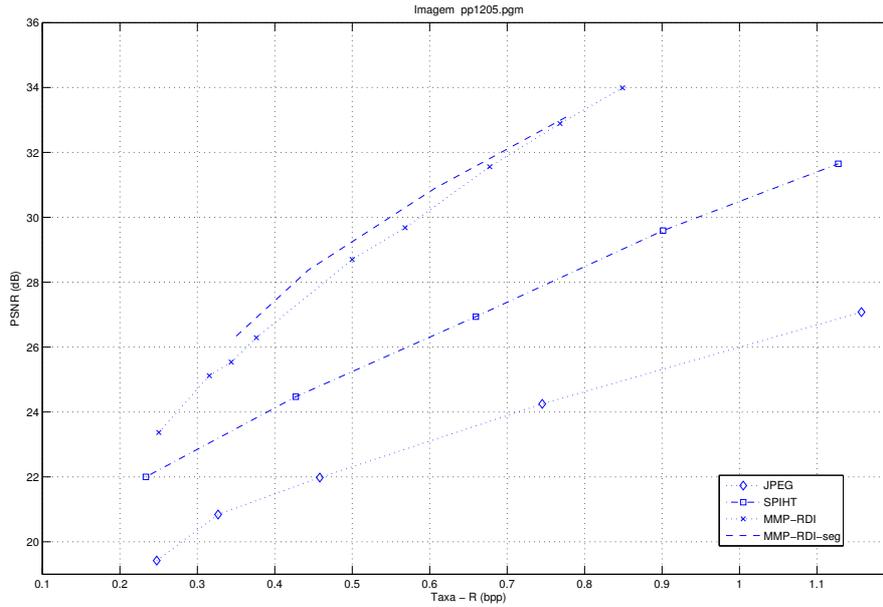


Figura 4.18: Curva Taxa-distorção do MMRDI, MMRDI-seg para imagem PP1205 (apêndice A.8) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

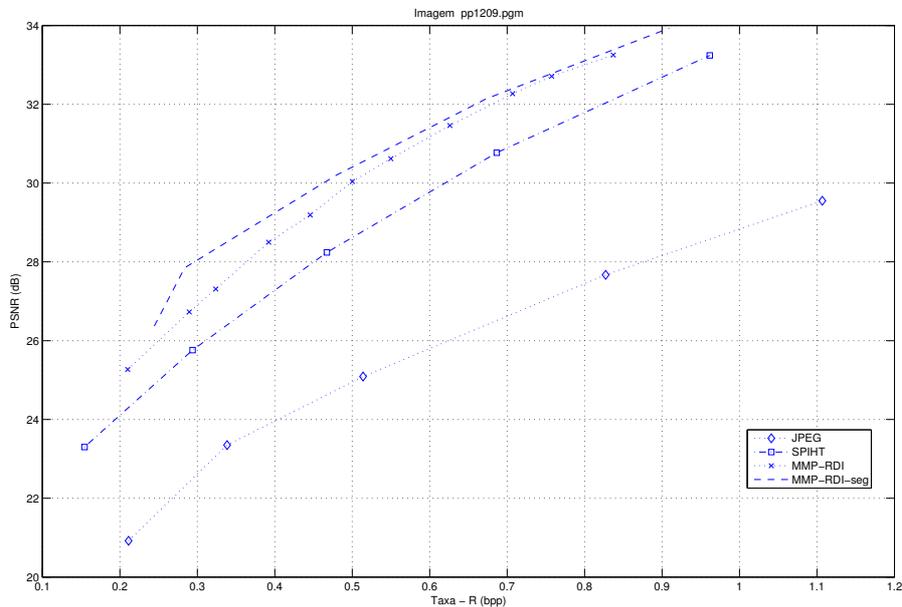


Figura 4.19: Curva Taxa-distorção do MMRDI, MMRDI-seg para imagem PP1209 (apêndice A.9) 512×512 comparando com o desempenho de outros algoritmos codificadores de imagens (SPIHT e JPEG).

Capítulo 5

MMPRDI usando blocos rotacionados e transpostos

Neste capítulo será feito a avaliação do desempenho taxa-distorção do MMPRDI usando a segmentação no espaço de códigos. Além disso, faremos o levantamento estatístico da utilização de blocos rotacionados e transpostos por subdicionário para verificar se seria efetivo usar a este tipo de segmentação. O algoritmo é bem parecido com o *MMP otimização taxa-distorção Intermediário*. A diferença está justamente no processamento da atualização do dicionário. Além da palavra resultante da concatenação dos dois nós filhos obtidos da árvore que minimiza o custo langrangeano (ver capítulo 4), o dicionário será também atualizado com blocos rotacionados e transpostos desta palavra. A seguir descrevemos o algoritmo, dando ênfase à atualização do dicionário, apresentando os resultados do levantamento estatístico obtidos em simulações computacionais.

5.1 Descrição

No momento da compressão de uma imagem, o MMPRDI procura uma árvore de segmentação ótima que minimize o custo langrangeano total (ver capítulo 4). Então, os blocos concatenados resultantes desta árvore de segmentação serão codificados e o dicionário terá cada uma de suas escalas atualizadas com a transformação de escala deles. Além disso, o dicionário também será atualizado com variações (rotações e transposta) destes blocos. A próxima seção explica sobre a atualização do dicionário.

5.2 Atualização do Dicionário

Vimos na seção anterior que o algoritmo MPRDI usando blocos rotacionados e transpostos tem seu dicionário atualizado tanto com a transformação de escala dos blocos concatenados resultantes da árvore de segmentação que minimiza o custo lagrangeano, como também pela transformação de escala de suas variações. Estas variações são as rotações de 90° , 180° e 270° , além da transposta (bloco transposto), destes blocos, como também rotações de 90° , 180° e 270° do bloco transposto. Por exemplo, suponhamos que o bloco concatenado resultante seja o bloco da figura 5.1(a). Este bloco passa por uma transformação de escala (sofre contrações ou dilatações) e atualiza cada escala do dicionário. Em seguida, este bloco é rotacionado de 90° (figura 5.1(b)). O bloco obtido desta rotação passa pelo mesmo processo que o bloco da figura 5.1(a). Ele é contraído ou dilatado (através da transformação de escala), e o resultado da transformação é usado para atualizar cada escala do dicionário. E este processo continua com as outras rotações de 180° e 270° (figuras 5.1(c) e 5.1(d)), a transposta do bloco concatenado (bloco transposto, figura 5.1(e)) e as suas rotações de 90° e 180° (figuras 5.1(f) e 5.1(g)), e finalizando com a rotação de 270° do bloco transposto. Assim, o dicionário terá 8 tipos de atualização para cada escala do dicionário.

5.3 Curvas taxa-distorção

Nesta seção apresentamos resultados experimentais de simulações no computador do MPRDI e do MPRDI com a atualização do dicionário incluindo blocos rotacionados e transpostos para uma série de imagens de teste. Mostraremos os resultados obtidos para as mesmas imagens de testes usadas no capítulo 4, todas de tamanho 512×512 . A imagem PP1205 (ver capítulo 4) é composta de texto e fórmulas (imagens mistas). A imagem PP1209 é composta de texto, gráficos e duas figuras da imagem Lena. Além dos resultados do MPRDI com a atualização do dicionário incluindo blocos rotacionados e transpostos, inserimos também nos gráficos, para efeito de comparação do desempenho entre eles, os resultados do algoritmo SPIHT [27, 28], que ilustra o desempenho dos codificadores de imagens modernos baseados em wavelets [10][11][12] e o JPEG [14]. As figuras 5.2 a 5.8

013	125	208	134
013	013	208	134
013	125	208	134
125	013	056	134

(a)

134	134	134	134
208	208	208	056
125	013	125	013
013	013	013	125

(b)

134	056	013	125
134	208	125	013
134	208	013	013
134	208	125	013

(c)

125	013	013	013
013	125	013	125
056	208	208	208
134	134	134	134

(d)

013	013	013	125
125	013	125	013
208	208	208	056
134	134	134	134

(e)

125	013	056	134
013	125	208	134
013	013	208	134
013	125	208	134

(f)

134	134	134	134
056	208	208	208
013	125	013	125
125	013	013	013

(g)

134	208	125	013
134	208	013	013
134	208	125	013
134	056	013	125

(h)

Figura 5.1: Exemplo de um bloco concatenado e suas rotações, transposta e rotações de sua transposta. (a) mostra o bloco concatenado para atualizar o dicionário e para realizar as rotações, além da transposta, e rotações do bloco transposto; (b), (c) e (d) São as rotações de 90° , 180° e 270° do bloco concatenado, respectivamente; (e) Transposta do bloco concatenado; (f), (g) e (h) Bloco transposto rotacionado de 90° , 180° e 270° , respectivamente.

mostram que o MPRDI com atualização do dicionário incluindo blocos rotacionados e transpostos teve um desempenho melhor que o MMMPRDI sem as modificações no dicionário (MPRDI normal). Podemos verificar que o desempenho deste MPRDI aproxima-se cada vez mais do desempenho do SPIHT. Quanto às figuras 5.9 e 5.10 podemos observar que o algoritmo MPRDI teve o desempenho melhor que o JPEG e SPIHT, conforme descrito no capítulo anterior. O mesmo aconteceu com o MPRDI com a atualização do dicionário incluindo blocos rotacionados e transpostos, tendo um desempenho melhor que o JPEG e o SPIHT, mas não superou o MPRDI normal para as imagens PP1205 e PP1209. Nota-se que este mesmo MPRDI para as imagens PP1205 e PP1209, teve o desempenho melhor que o SPIHT. Isto quer dizer que, mesmo com o MPRDI tendo seu dicionário atualizado de formas diversas, ela codifica as imagens que são constituídas de textos, figuras e equações (imagens mistas), com um desempenho melhor que o dos outros algoritmos. Podemos ver pelos gráficos das figuras 5.2 a 5.10 que mesmo aumen-

tando o tamanho do dicionário com blocos rotacionados e transpostos, o MMPRDI teve uma discreta melhora na relação sinal ruído em algumas imagens e piora em outras. As imagens que apresentaram melhoras na relação sinal ruído foram a aerial, baboon, bridge, e a lena. O aumento de desempenho obtido foi, em geral, de pouco menos de 0.5dB a 0.5bpp.

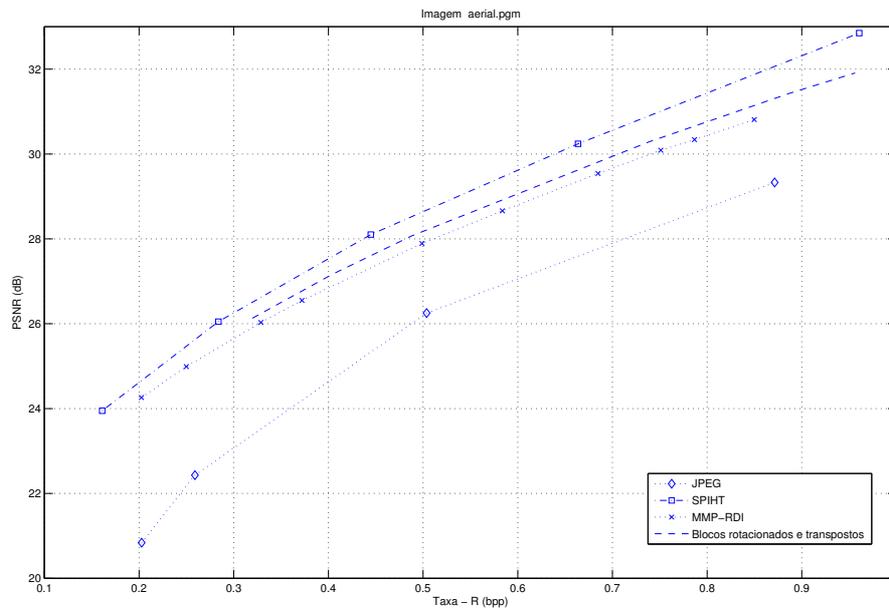


Figura 5.2: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Aerial 512×512 .

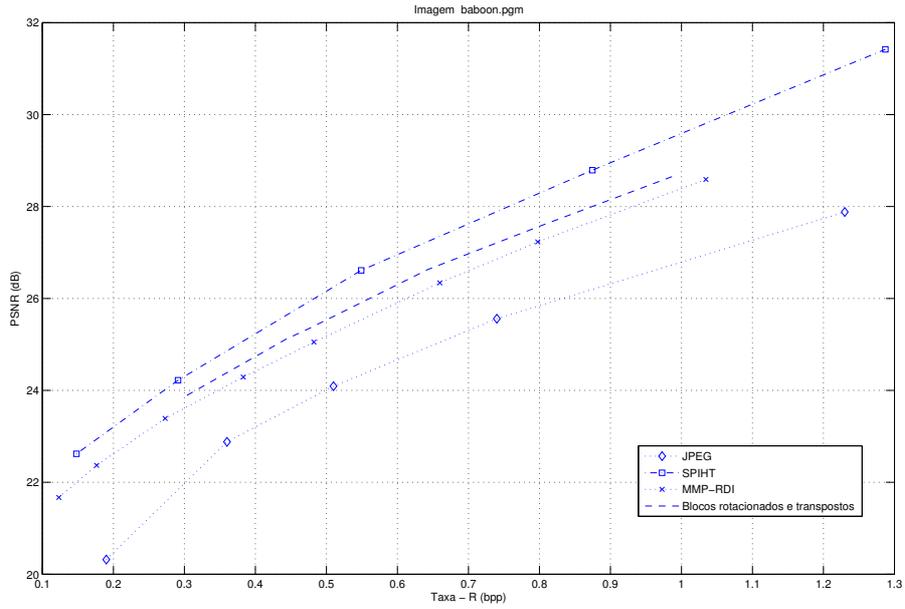


Figura 5.3: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Baboon 512×512 .

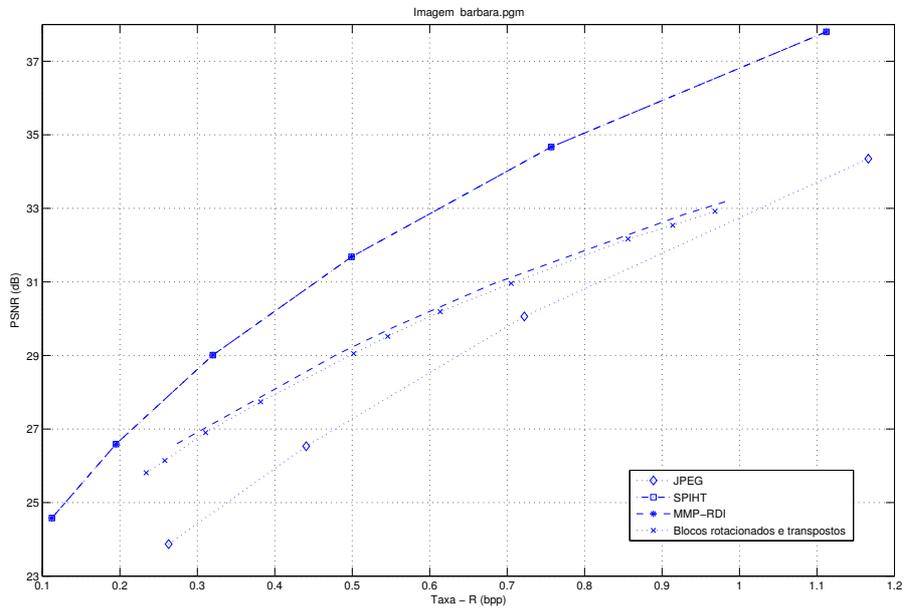


Figura 5.4: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Barbara 512×512 .

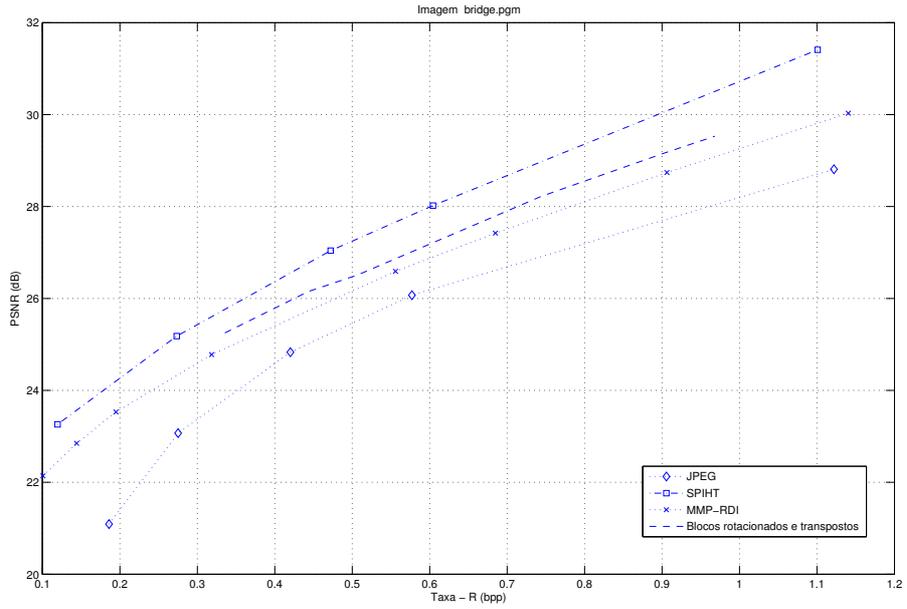


Figura 5.5: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Bridge 512×512 .

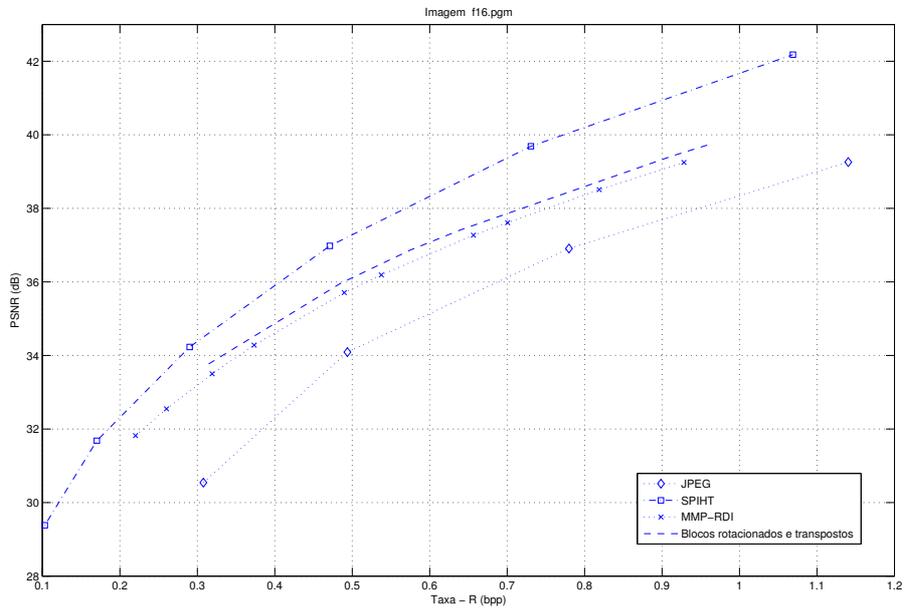


Figura 5.6: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem F16 512×512 .

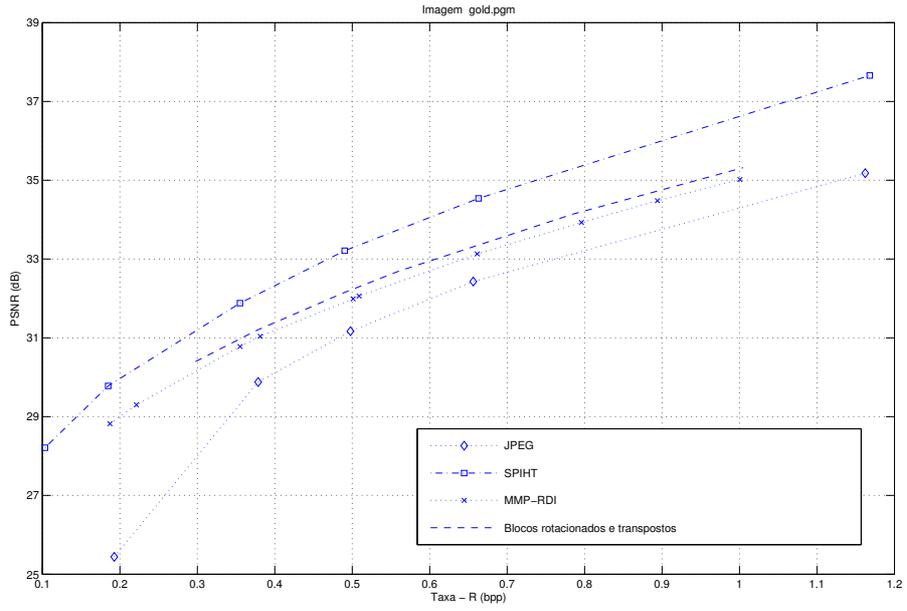


Figura 5.7: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Gold 512×512 .

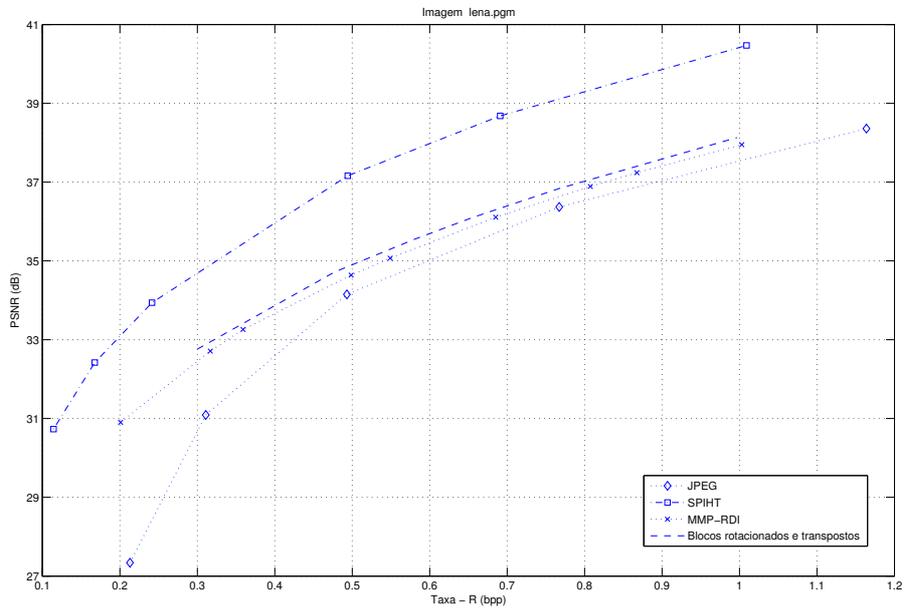


Figura 5.8: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem Lena 512×512 .

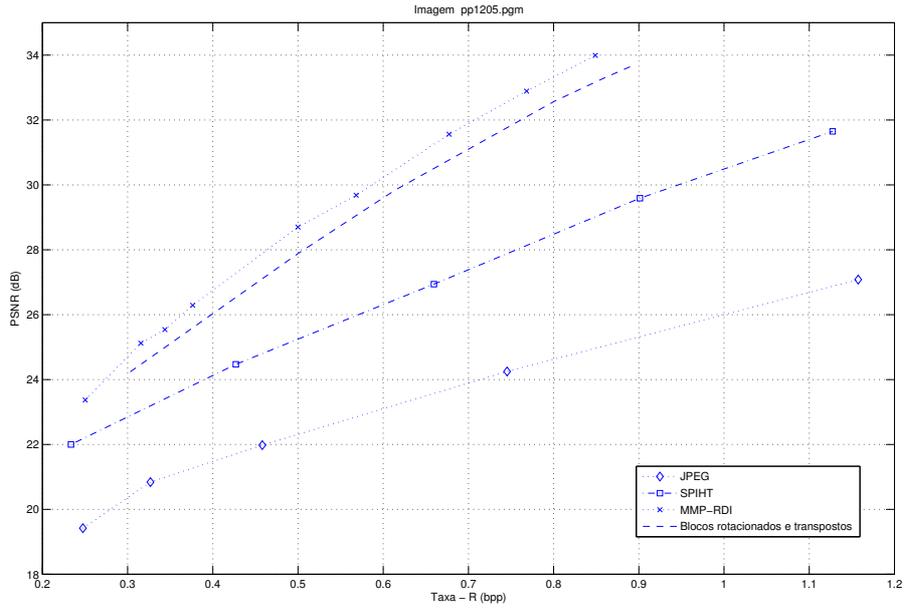


Figura 5.9: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem PP1205 512×512 .

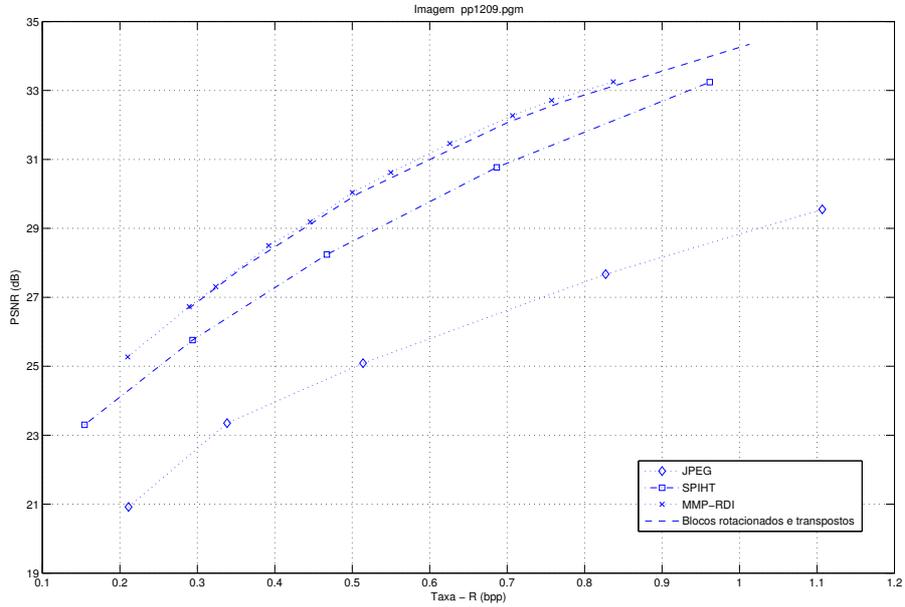


Figura 5.10: Curva Taxa-distorção do MMPRDI, e blocos rotacionados e transpostos para imagem PP1209 512×512 .

5.4 Resultados

Nesta seção mostramos os resultados do levantamento estatístico das probabilidades das frequências dos vetores dos subdicionários do MMPRDI usando blocos rotacionados e transpostos, a fim de verificar se são bastante diferentes entre si, ou seja, avaliar se seria efetivo usar a segmentação no espaço de códigos. As figuras 5.11 a 5.19 mostram os resultados da simulação do algoritmo com as imagens de teste do apêndice A), as mesmas usadas no capítulo 4. A apresentação dos resultados segue o mesmo padrão adotado no capítulo 4. Podemos perceber pelas figuras 5.12 a 5.15 e as figuras 5.17 a 5.19 que as escalas mais usadas para compor cada escala do dicionário continuam sendo as escalas com blocos de tamanho 4×8 . Somente a figura 5.11 mostrou um resultado diferente das outras; nesta, a escala que teve maior participação na composição do dicionário foi a de tamanho 8×8 . As figuras 5.18 e 5.19 mostram que a escala 1×1 do dicionário não foi totalmente constituída pela escala 4×4 , e sim pela 2×4 . Os resultados estatísticos obtidos mostram quais escalas do dicionário foram mais utilizadas para atualizá-lo. Na próxima seção descrevemos os resultados relativos a quais tipos de variações do bloco concatenado (bloco resultante da árvore de segmentação que minimiza o custo lagrangeano, ver seção 5.2) que foram mais usados.

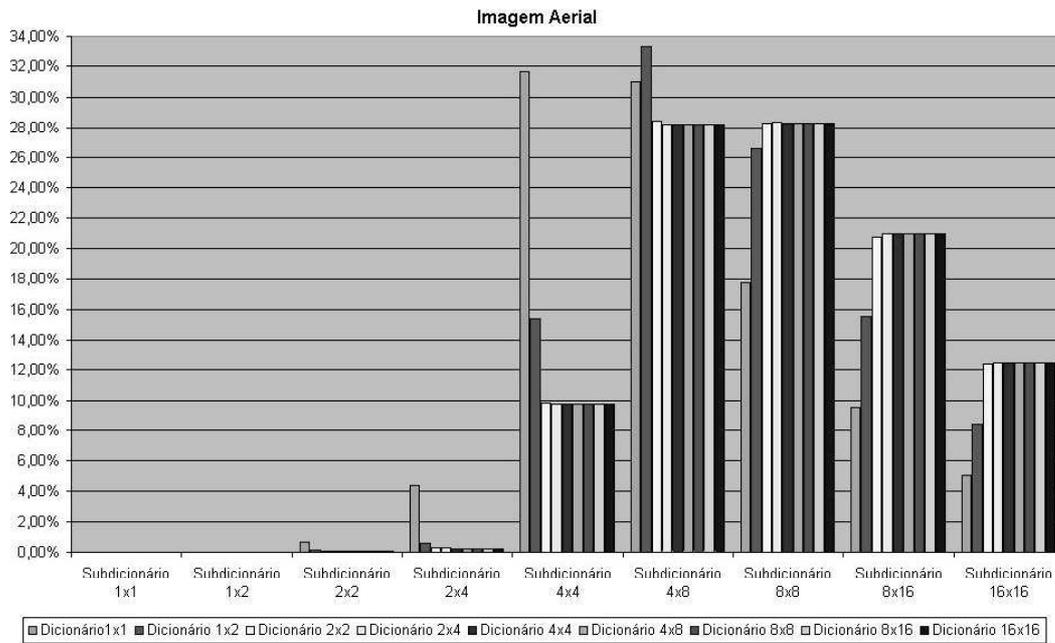


Figura 5.11: Resultado do levantamento estatístico do MPPRDI para a imagem de teste aerial com rotação e transposta de blocos.

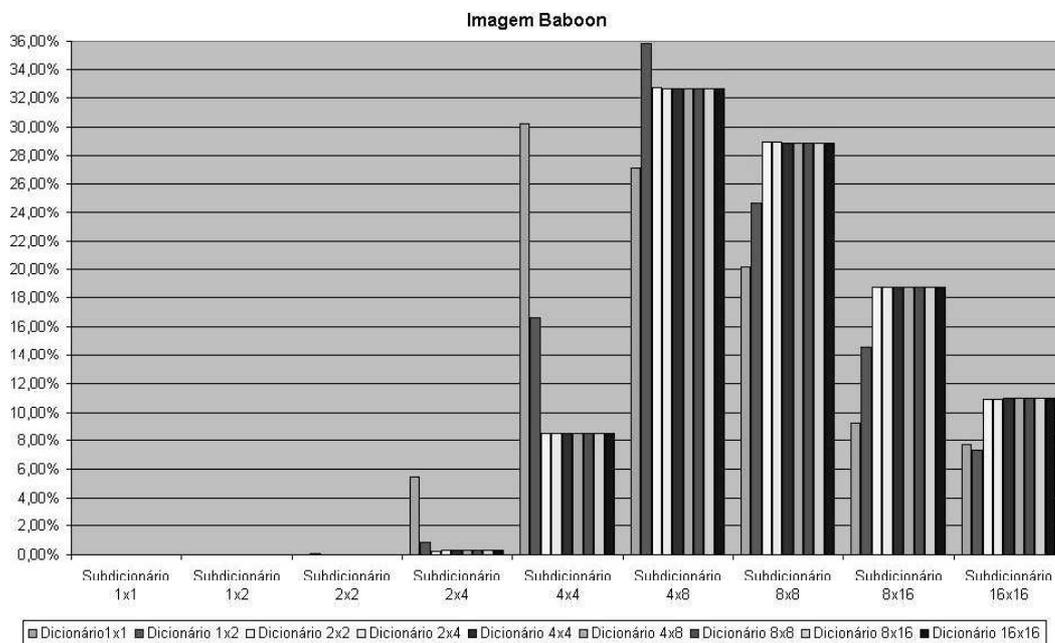


Figura 5.12: Resultado do levantamento estatístico do MPPRDI para a imagem de teste baboon com rotação e transposta de blocos.

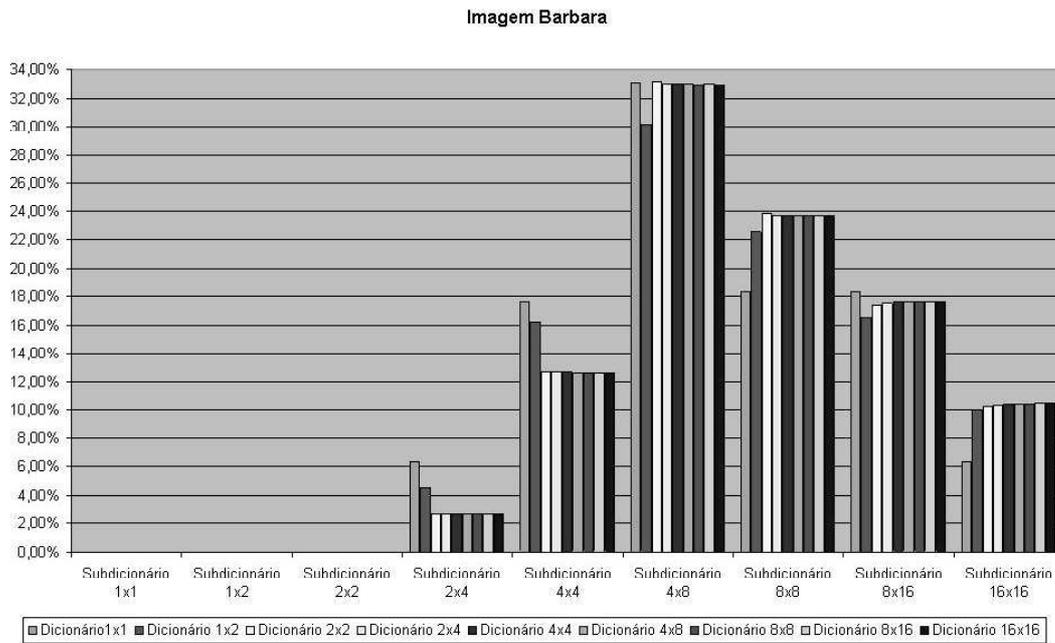


Figura 5.13: Resultado do levantamento estatístico do MPRDI para a imagem de teste barbara com rotação e transposta de blocos.

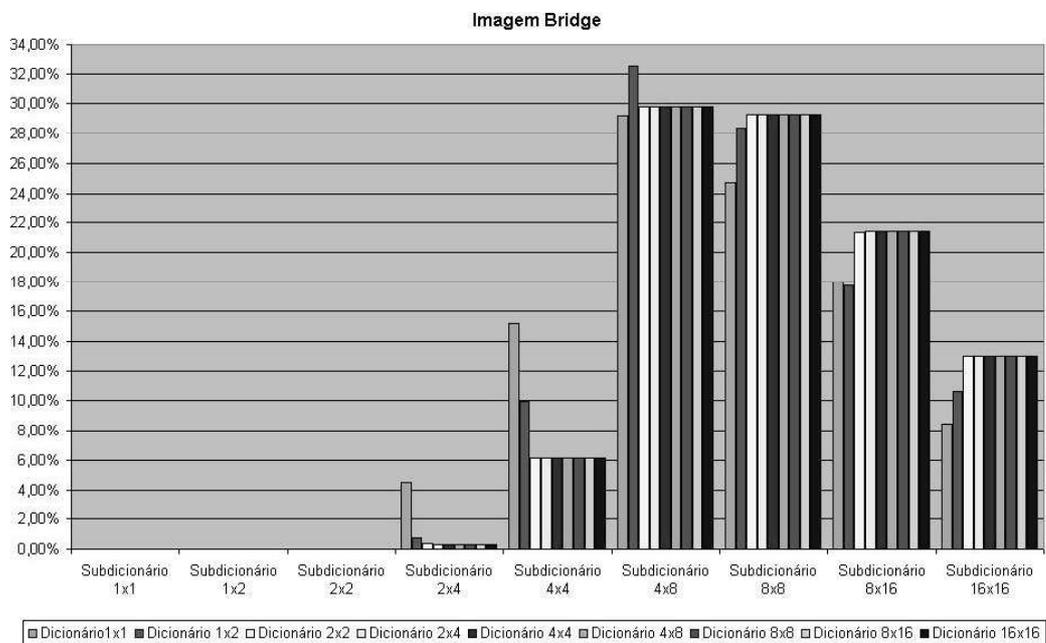


Figura 5.14: Resultado do levantamento estatístico do MPRDI para a imagem de teste bridge com rotação e transposta de blocos.

Imagem F16

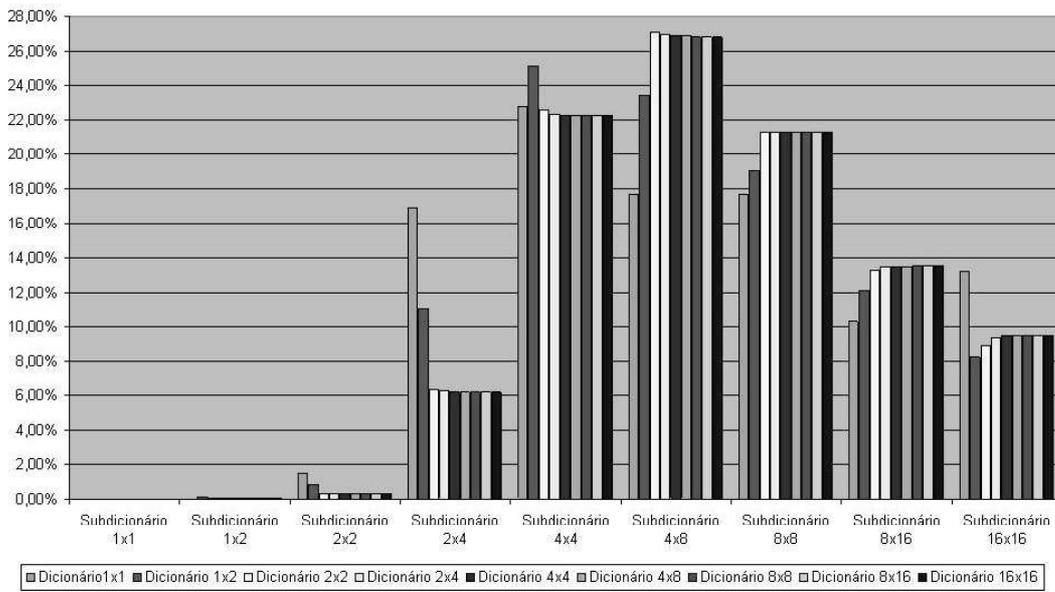


Figura 5.15: Resultado do levantamento estatístico do MPRDI para a imagem de teste f16 com rotação e transposta de blocos.

Imagem Gold

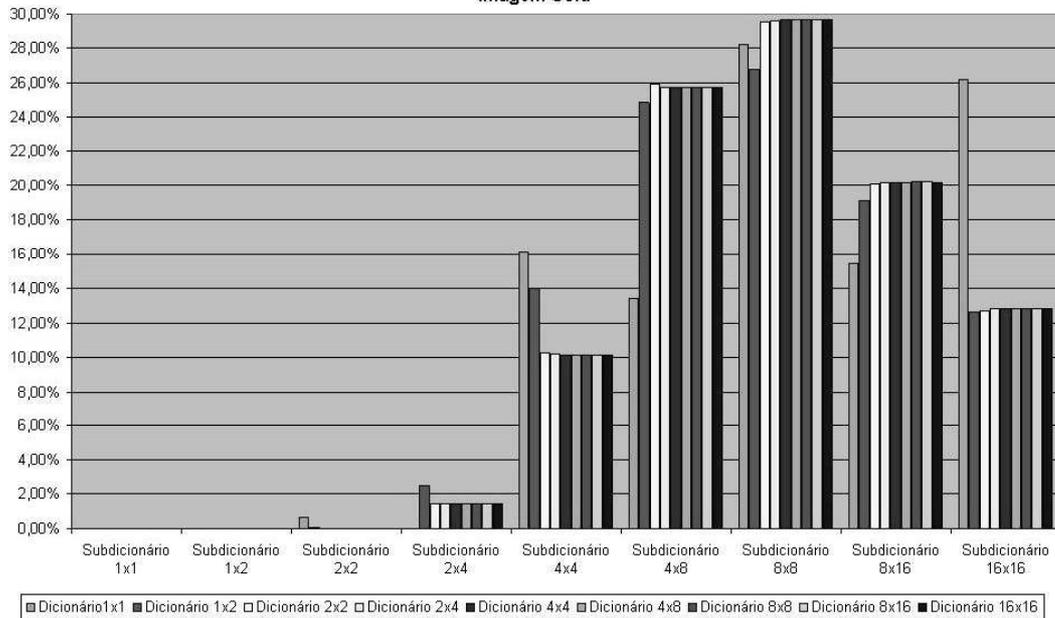


Figura 5.16: Resultado do levantamento estatístico do MPRDI para a imagem de teste gold com rotação e transposta de blocos.

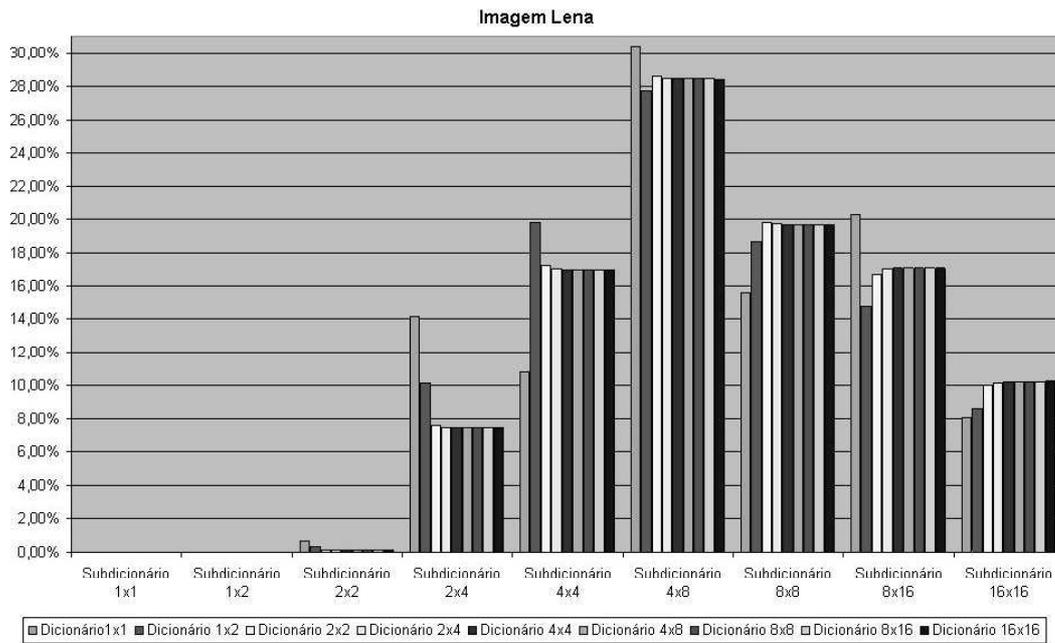


Figura 5.17: Resultado do levantamento estatístico do MPRDI para a imagem de teste lena com rotação e transposta de blocos.

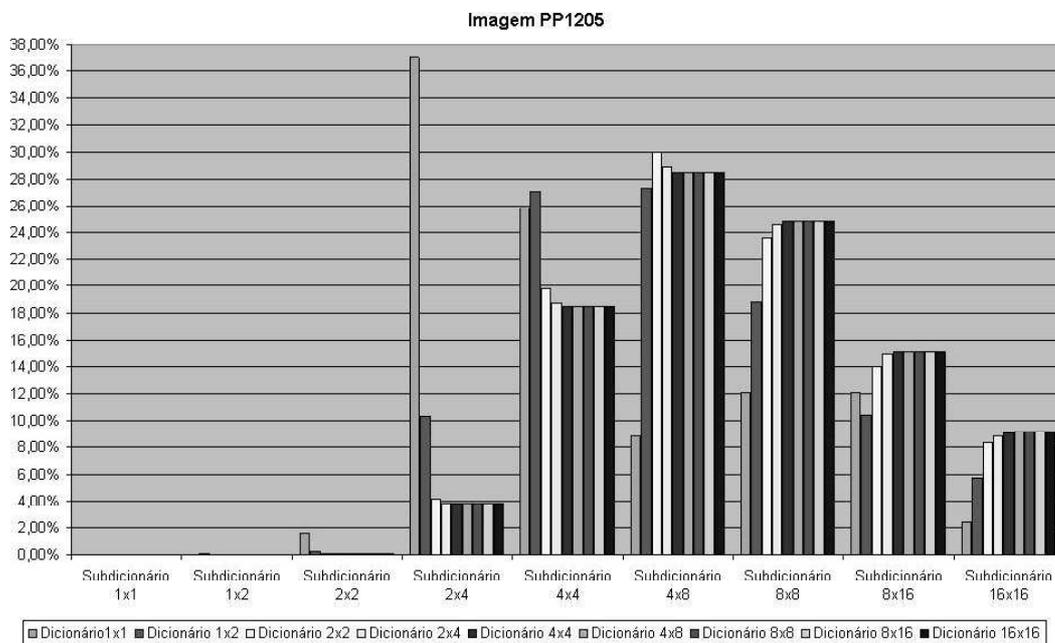


Figura 5.18: Resultado do levantamento estatístico do MPRDI para a imagem de teste pp1205 com rotação e transposta de blocos.

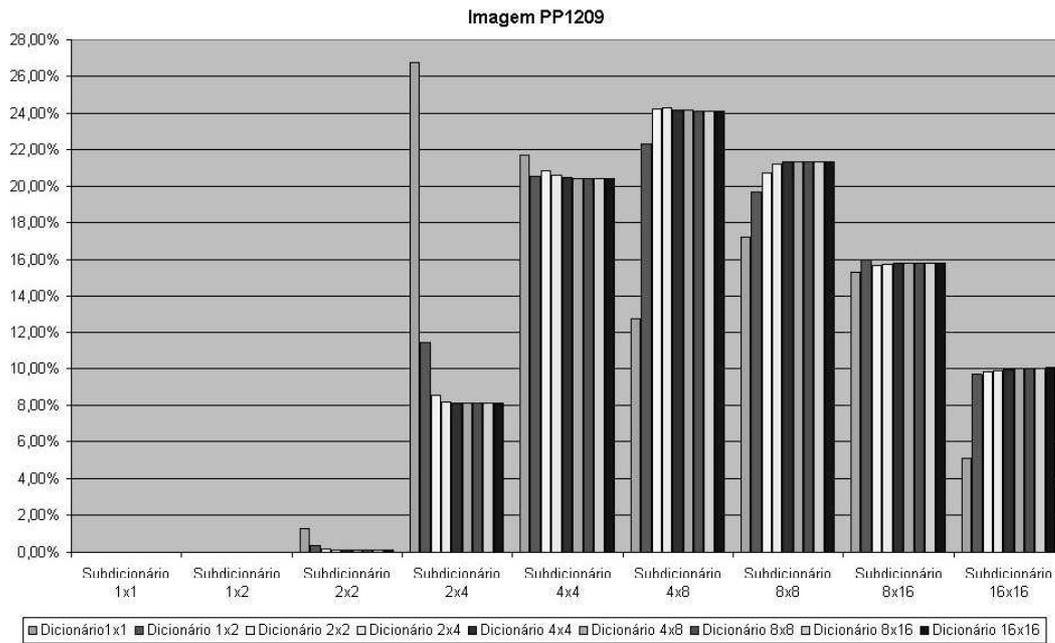


Figura 5.19: Resultado do levantamento estatístico do MPRDI para a imagem de teste pp1209 com rotação e transposta de blocos.

5.4.1 Resultados dos blocos rotações e transpostas mais usados

Sabemos que a escala que mais contribuiu para construção de cada escala do dicionário foi a 4×8 . Como o dicionário é atualizado com o bloco concatenado e suas variações (blocos rotacionados e transpostos, ver seção 5.2), fica interessante descobirmos quais destas variações mais contribuíram para a sua atualização. Esta seção mostra justamente os resultados deste levantamento estatístico. Podemos observar pelas figuras 5.20 a 5.28 que não há muita diferença entre as mesmas para praticamente todas as escalas. Os gráficos mostram que as probabilidades dos vetores terem sido gerados a partir das diversas variações, estão muito próximas, o que não torna útil o uso da segmentação no espaço de códigos (ver capítulo 4 seção 4.4) levando em conta a codificação do flag informando qual o tipo de bloco que está sendo usado, bloco concatenado e/ou blocos rotacionados e transpostos, e um índice indicando a sua posição dentro de cada variação (rotações de 90° , 180° e 270° , e transposições). Porém, pelos resultados dos gráficos nas figuras 5.11 a 5.19 pode-

se observar que existe uma chance de ser útil o uso da segmentação no espaço de códigos, com um flag informando de qual subdicionário pertence o bloco concatenado e um índice indicando sua localização dentro do subdicionário (conforme mostrado no capítulo 4), mas devido as restrições de tempo não foi possível implementar este algoritmo.

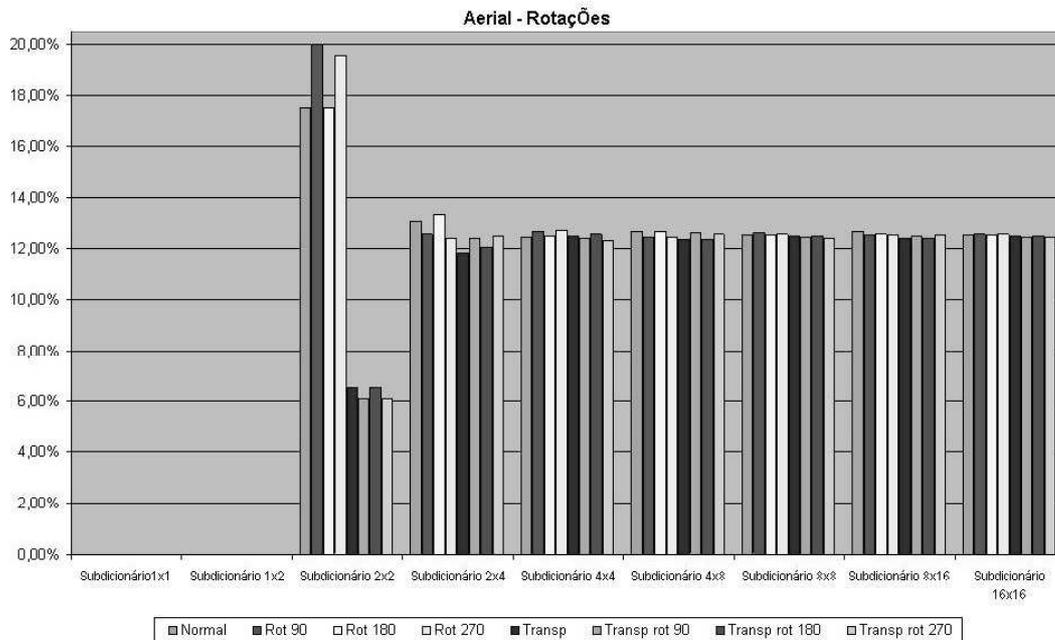


Figura 5.20: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste aerial.

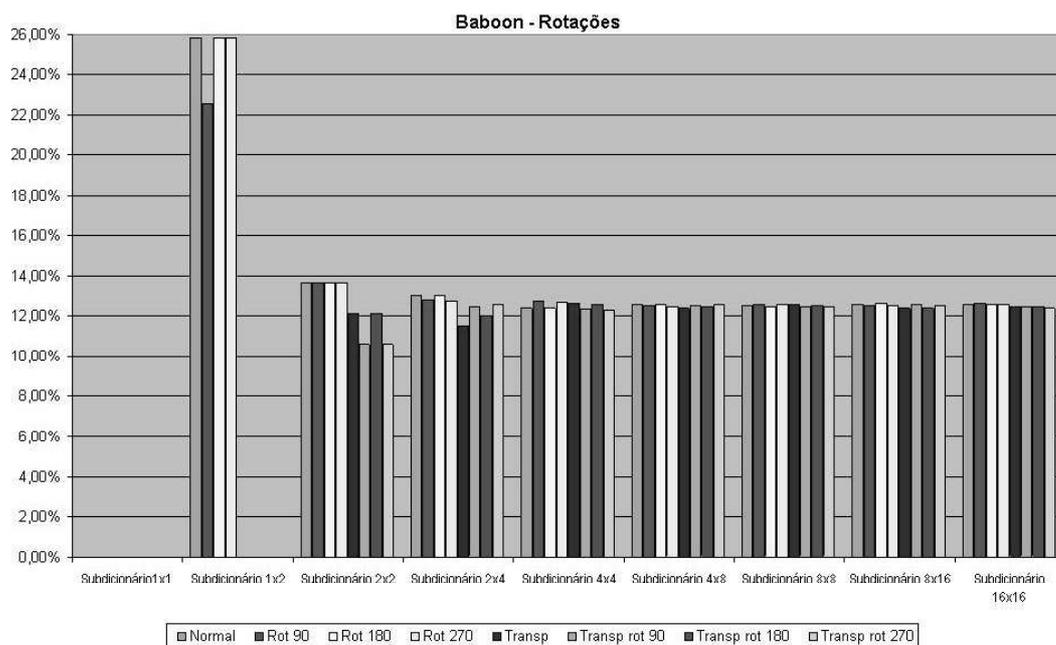


Figura 5.21: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste baboon.

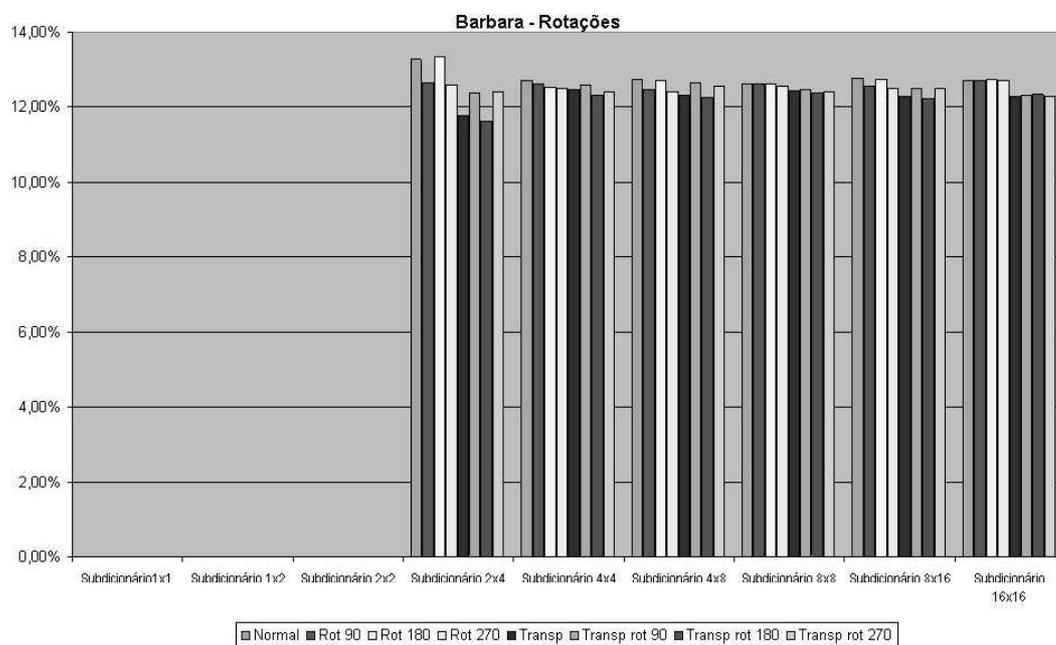


Figura 5.22: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste barbara.

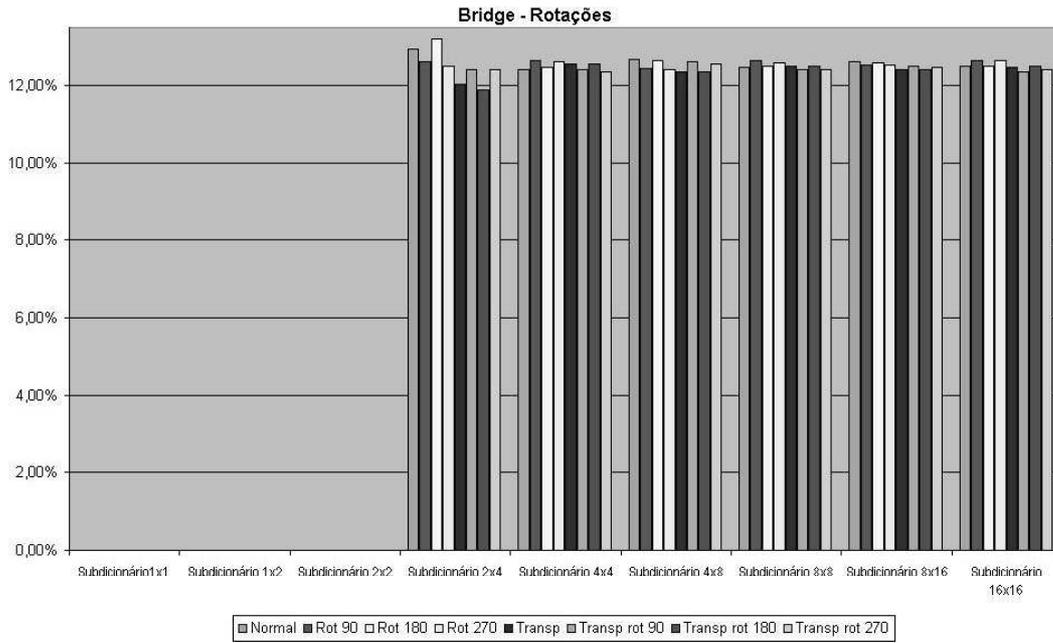


Figura 5.23: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste bridge.

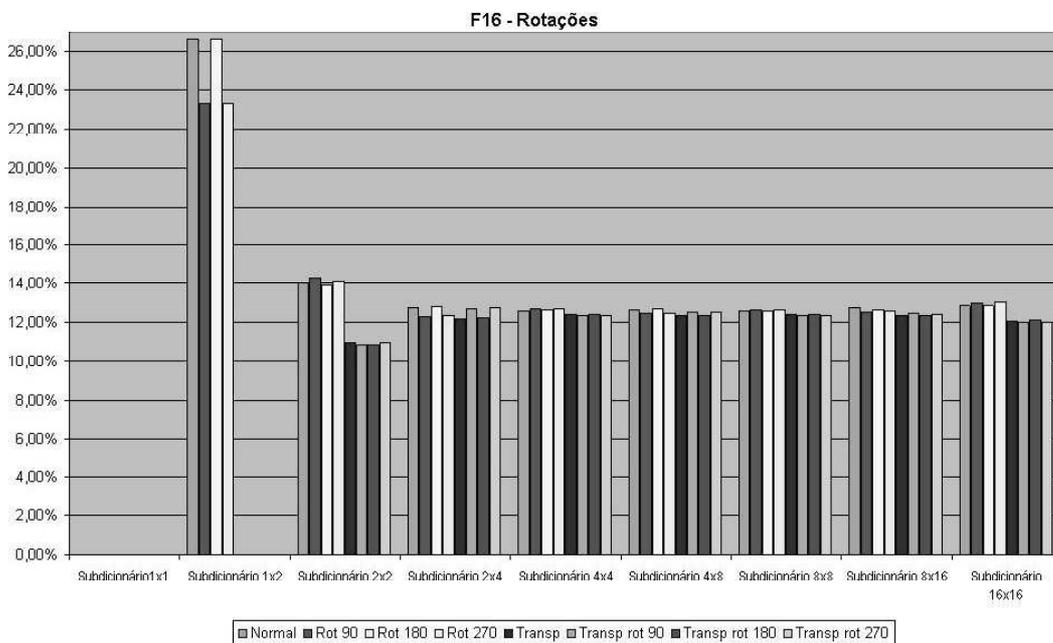


Figura 5.24: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste f16.

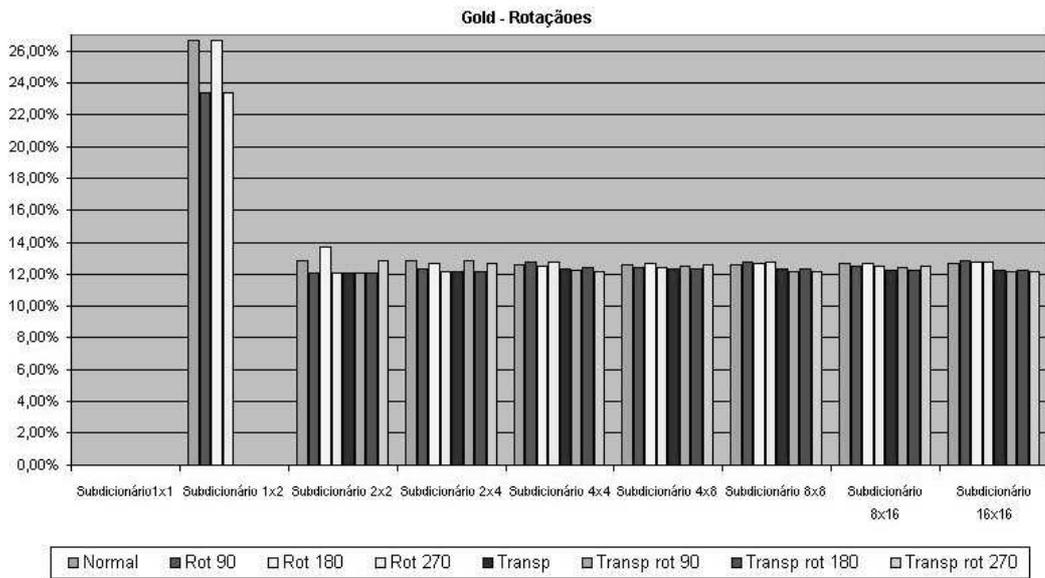


Figura 5.25: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste gold.

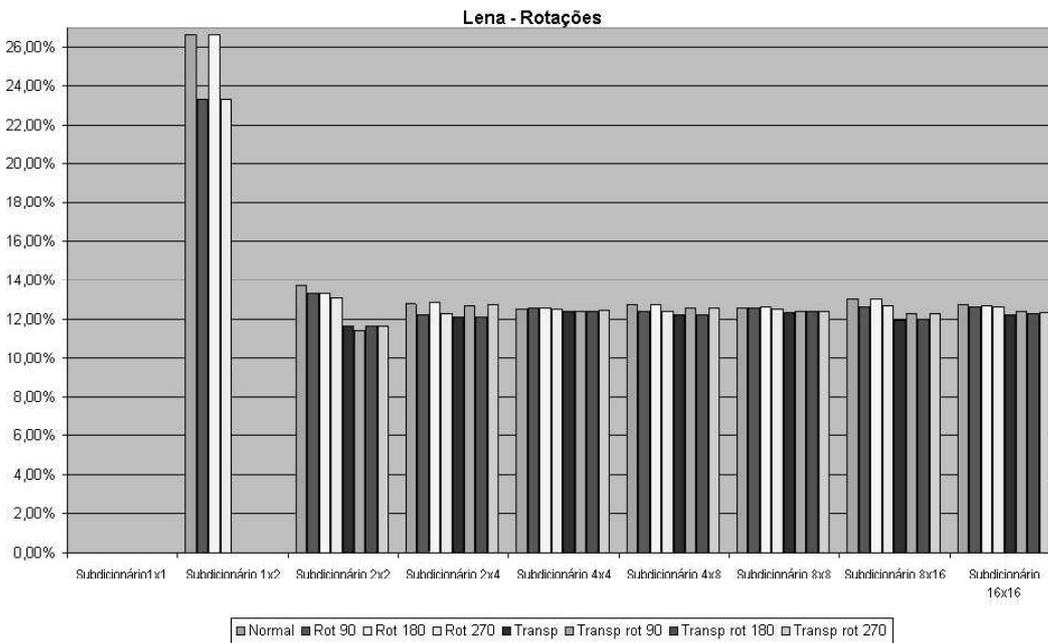


Figura 5.26: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste lena.

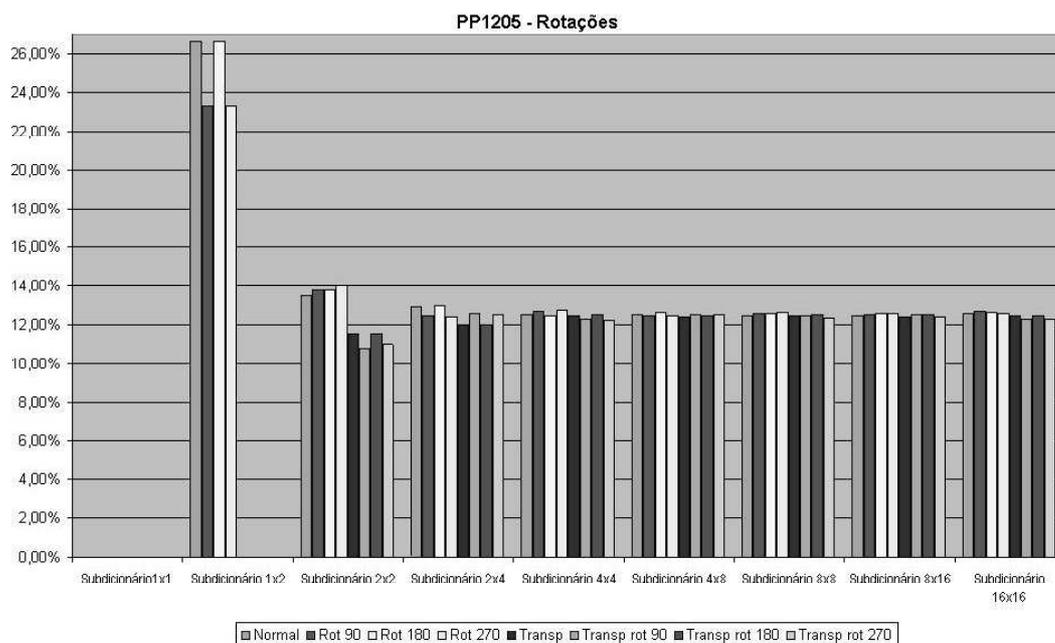


Figura 5.27: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste pp1205.

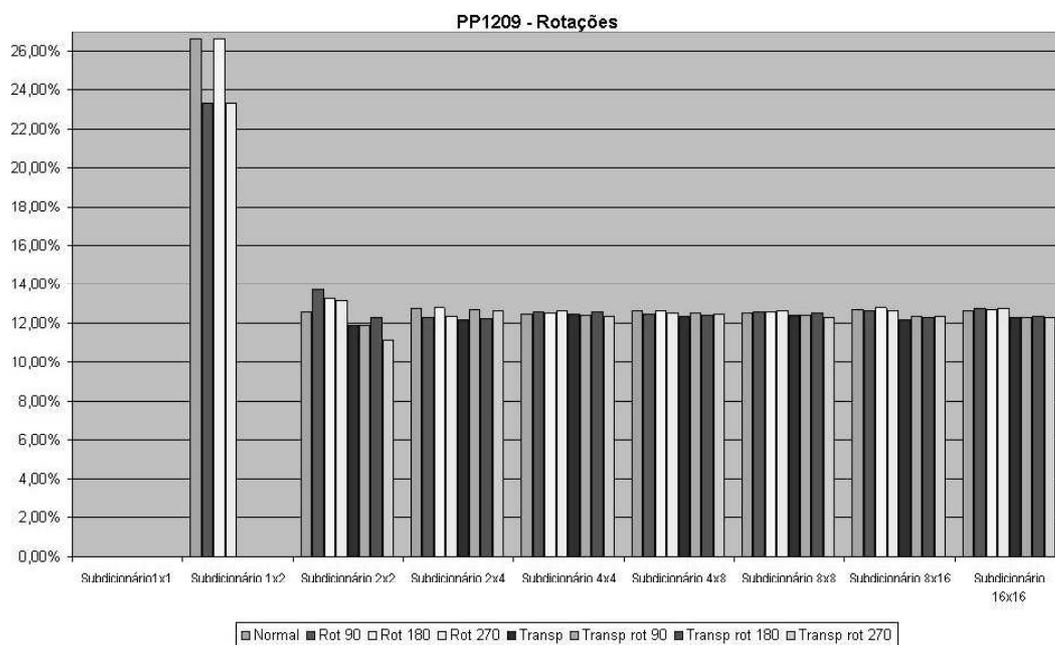


Figura 5.28: Resultado mostrando as rotações e transposta de blocos mais usados para atualizar o dicionário com a imagem de teste pp1209.

Capítulo 6

MMPRDI com deslocamento de blocos

No capítulo anterior a proposta de modificação no algoritmo MMPRDI consistia em incluir no dicionário, no momento de sua atualização, blocos rotacionados e transpostos. O dicionário era atualizado com as rotações e transposições, de um bloco concatenado (ver capítulo 5). Agora, o dicionário será atualizado com deslocamento de blocos de tamanho 1×1 , 1×2 , e 2×2 , além do tamanho 2×4 , 4×4 , e do 4×8 , como também o 8×8 , 8×16 e o 16×16 , deslocados de múltiplos $1/2$, $1/4$ e $1/8$ do seu tamanho em cada direção. Veremos, neste capítulo, o levantamento estatístico das probabilidades das freqüências dos vetores dos subdicionários com deslocamento de blocos para verificar se é efetivo ou não o uso da segmentação no espaço de códigos incluindo no dicionário blocos deslocados de múltiplos $1/2$, $1/4$ e $1/8$ do seu tamanho. As próximas seções descrevem o algoritmo, com ênfase na atualização do dicionário e mostram os resultados obtidos nas simulações.

6.1 Descrição

A descrição é similar à do capítulo 5. A diferença está na atualização do dicionário. O dicionário agora é atualizado com deslocamento de blocos de tamanhos determinados, deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho sobre os blocos codificados da imagem de entrada. Após a codificação do bloco, também são incluídos no dicionário blocos resultantes do janelamento da imagem reconstruída

até então. A janela é deslocada dos múltiplos $1/2$, $1/4$ e $1/8$ do seu tamanho em cada direção (na verdade para cima e para esquerda da imagem reconstruída).

6.2 Atualização do Dicionário

No capítulo 4 cada escala do dicionário era atualizada, após a codificação, com a transformação de escala do bloco concatenado. No capítulo 5 além desta atualização também tínhamos a que envolvia variações (rotações e transposições) deste bloco. Neste capítulo a atualização do dicionário inclui tanto a transformação de escala do bloco concatenado quanto os deslocamentos de blocos de tamanhos determinados sobre a imagem reconstruída, de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho, ou seja, os blocos são obtidos a partir de janelas de tamanhos determinados que “passam” sobre partes da imagem que já foram codificadas. A figura 6.1 mostra como será feita a atualização do dicionário. Nesta figura, uma janela de tamanho 8×8 é deslocada pela parte da imagem que já foi codificada, sendo este deslocamento de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho. Em cada posição desta janela temos um bloco que será usado para atualizar o dicionário. Então ele passa por uma transformação de escala para atualizar cada escala do dicionário. Espera-se que assim haja uma melhora nos resultados das imagens a serem codificadas.

6.3 Curva taxa-distorção

Nesta seção apresentamos resultados de simulações do MMRDI usando uma série de imagens de teste como entrada, as mesmas usadas nas seções anteriores. Mostraremos os resultados das simulações com MMRDI contendo a atualização do dicionário incluindo blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho. As figuras 6.2 a 6.7 contém as curvas taxa-distorção para estas imagens. Inserimos para comparação de resultados, o desempenho de algoritmos codificadores de imagens como o SPIHT e JPEG. Podemos verificar pelas figuras 6.6 e 6.7 que os resultados da curva taxa-distorção para o MMRDI incluindo blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ de seu tamanho tiveram uma melhoria, principalmente quando usamos os blocos deslocados de múltiplos de $1/4$ de seu tamanho. Este deslocamento por sua vez, melhorou bastante o desempenho do MMRDI. Isto

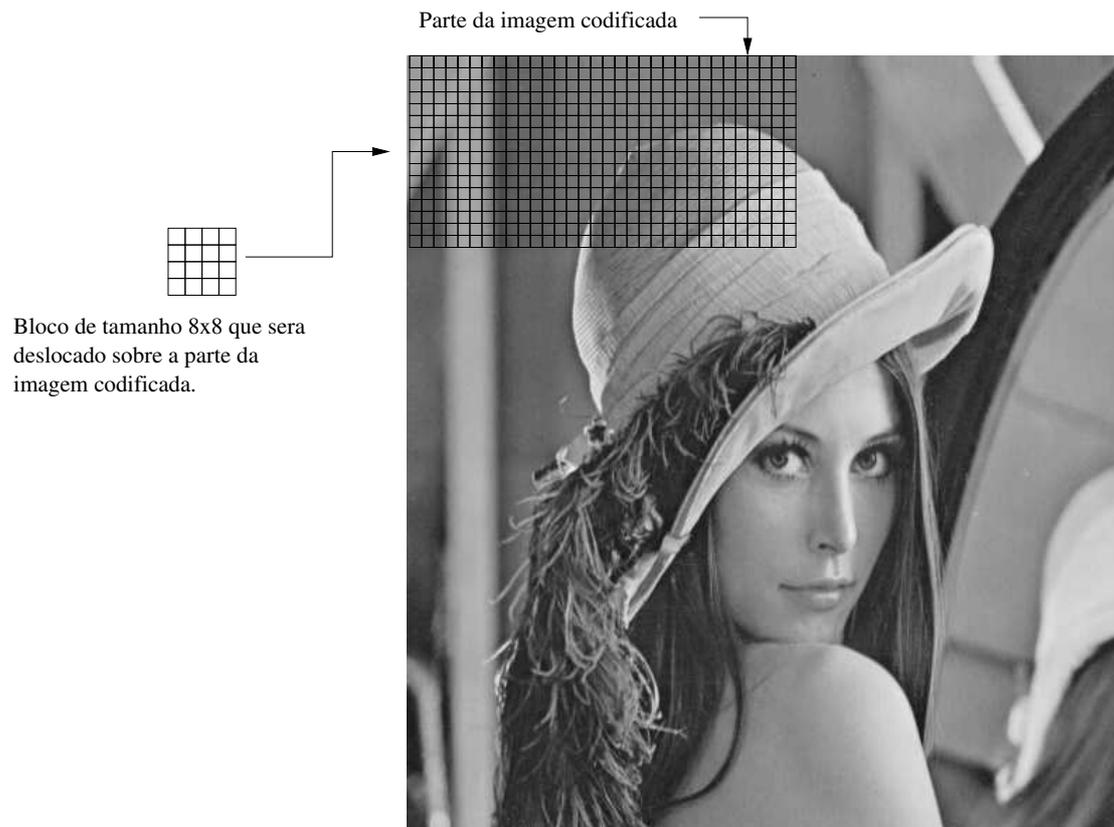


Figura 6.1: Exemplo de como será feito o deslocamento de blocos com comprimentos determinados. Neste exemplo o bloco utilizado tem tamanho 8×8

nos mostra que mesmo aumentando drasticamente a cardinalidade do dicionário, o atualizando com blocos deslocados, temos um ótimo resultado para imagens contendo texto e figuras. Para as figuras 6.2 a 6.5, o deslocamento de blocos de $1/4$ do tamanho ainda se destacou em relação aos outros deslocamentos, isto quer dizer que, quando atualizamos o dicionário com blocos deslocados, o melhor compromisso entre cardinalidade e desempenho taxa-distorção é obtido com deslocamentos de $1/4$ de seu tamanho.

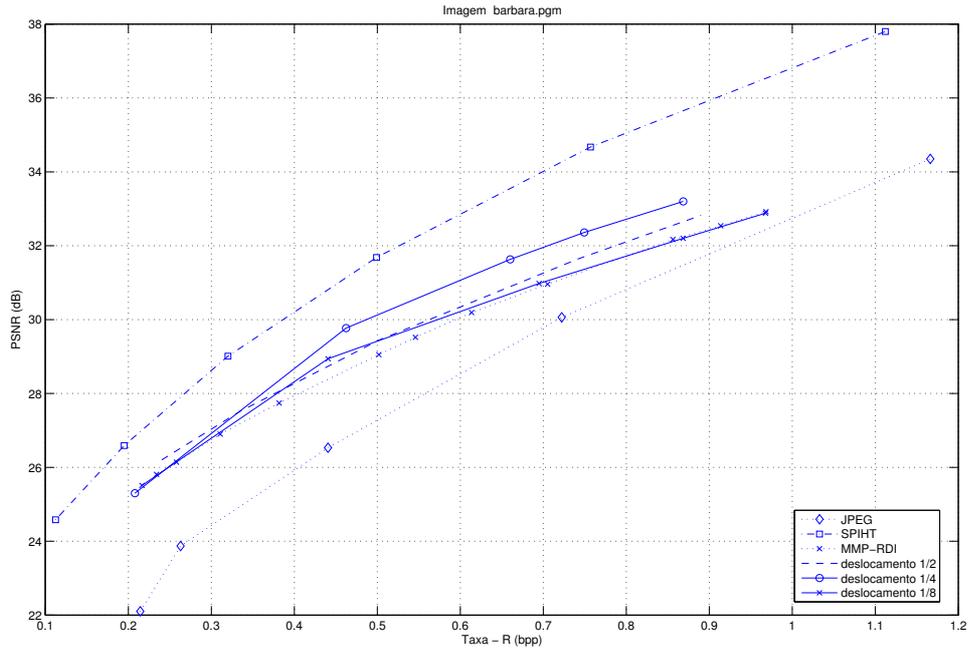


Figura 6.2: Curva Taxa-distorsão do MMRP-RDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem Barbara 512 × 512.

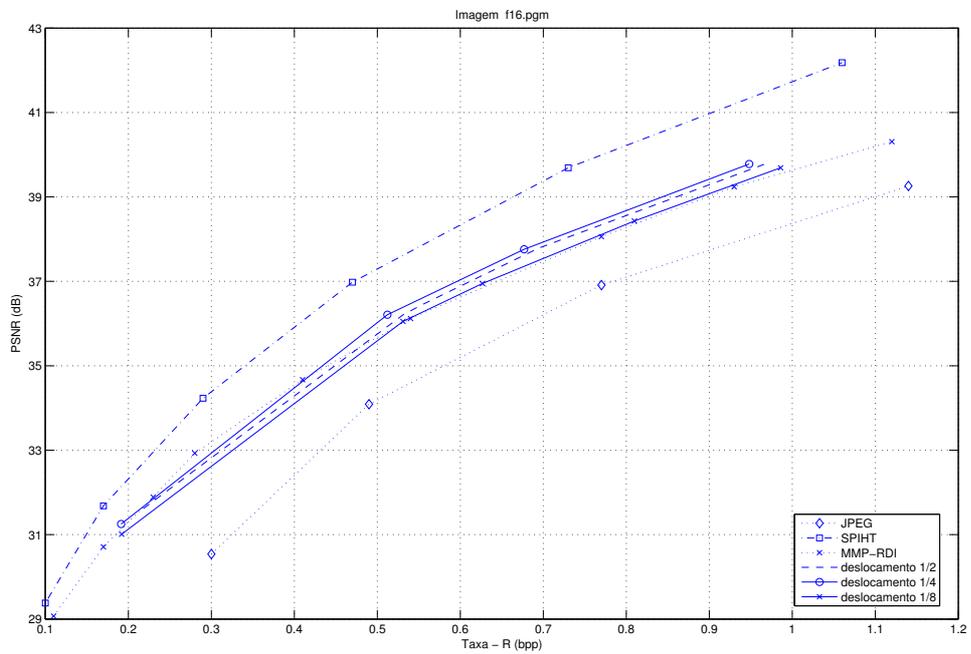


Figura 6.3: Curva Taxa-distorsão do MMRP-RDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem F16 512 × 512.

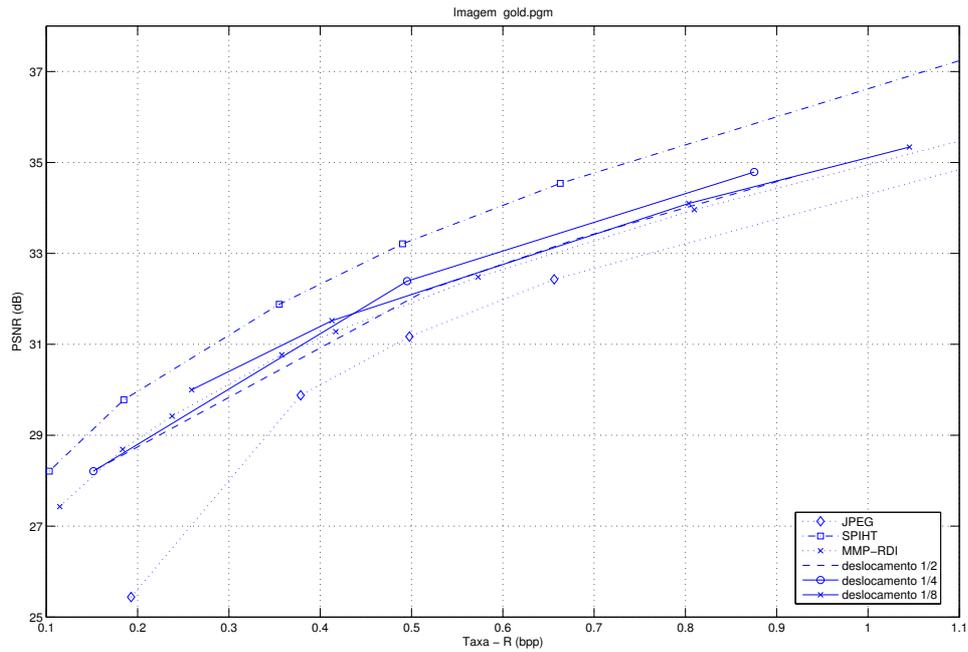


Figura 6.4: Curva Taxa-distorsão do MMRPDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem Gold 512×512 .

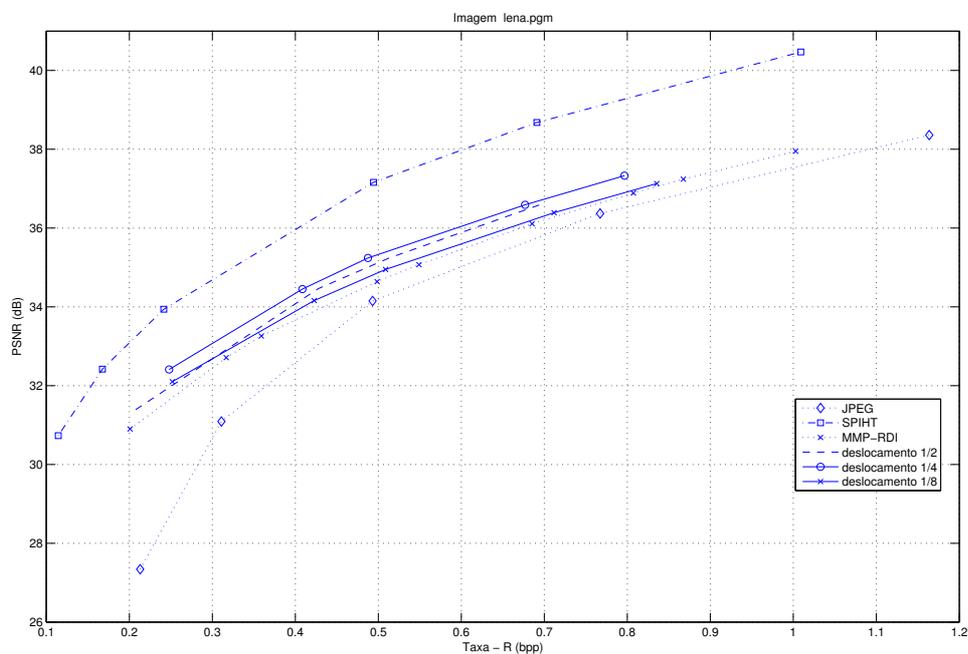


Figura 6.5: Curva Taxa-distorsão do MMRPDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem Lena 512×512 .

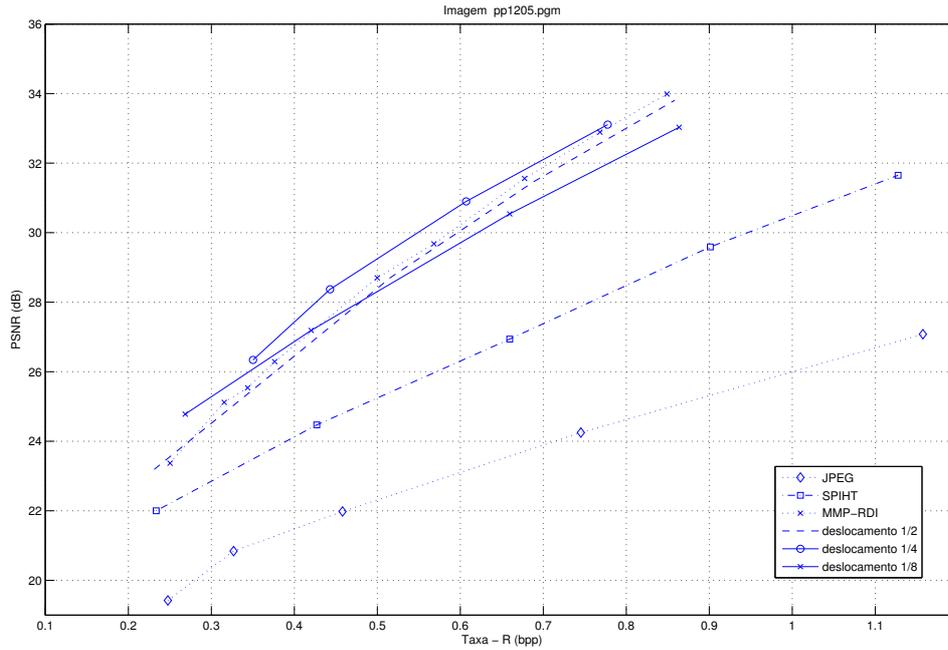


Figura 6.6: Curva Taxa-distorsão do MMPRDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem PP1205 512 × 512.

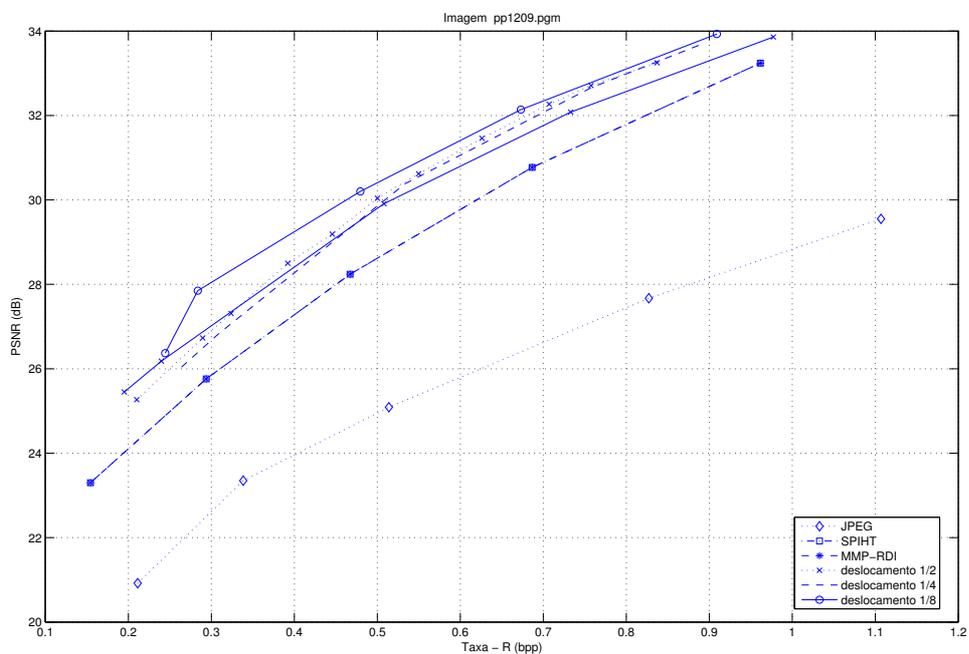


Figura 6.7: Curva Taxa-distorsão do MMP, MMPRDI, e blocos deslocados de múltiplos de 1/2, 1/4 e 1/8 do seu tamanho para imagem PP1209 512 × 512.

6.4 Resultados

Aqui mostramos os resultados da observação do comportamento da variação do MMP testada na seção anterior. Para facilitar esta análise cada subseção mostra os resultados para deslocamento de pixels mencionados anteriormente. Considerando que apenas seriam usados blocos com deslocamento múltiplos ou de $1/2$, ou de $1/4$ ou de $1/8$ de seu tamanho.

6.4.1 Resultados para adição de blocos deslocados de múltiplos de $1/2$ de seu tamanho

Nesta seção mostramos o resultado da medição das probabilidades dos vetores do dicionário do MMPRDI incluindo deslocamento de blocos de múltiplos de $1/2$ de seu tamanho. As figuras 6.8 a 6.13 mostram os gráficos do levantamento estatístico com as probabilidades dos vetores dos subdicionários. Nas abcissas dos gráficos temos os subdicionários, que informam de qual escala do dicionário veio o bloco concatenado contraído ou dilatado (transformação de escala); as ordenadas mostram as probabilidades de cada escala do dicionário. A soma de cada uma delas resulta em 100%; cada barra vertical correspondendo a uma determinada escala do dicionário (como mostra as legendas dos gráficos nas figuras 6.8 a 6.13). As figuras mostram que os dicionários mais utilizados para construir as outras escalas são os dicionários 2×2 , seguido do dicionário 2×4 , os quais ficaram bem próximos. Podemos observar também que os dicionários 1×1 e 1×2 foram mais atualizados com a escala 2×4 . Mas nota-se que na maior parte das imagens de teste o dicionário 2×2 foi o mais utilizado para construir as outras escalas do dicionário. A seção seguinte mostra quais foram as dimensões de blocos mais usadas para as mesmas imagens utilizadas nas simulações anteriores.

Imagem Barbara - com deslocamento de 1/2 comprimento do bloco

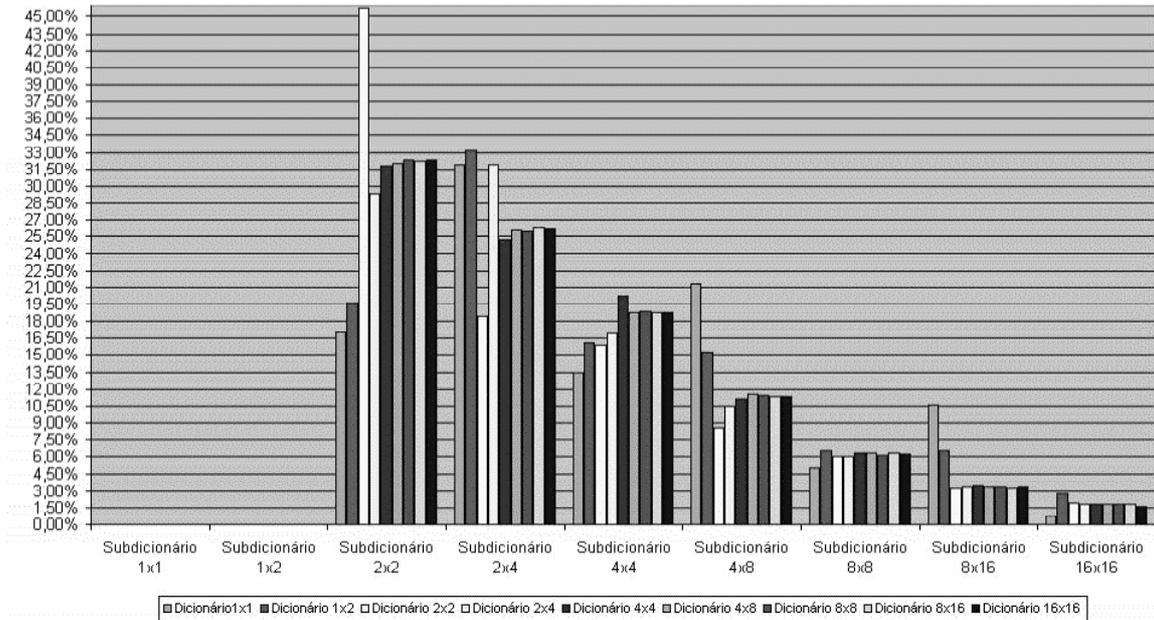


Figura 6.8: Resultado do levantamento estatístico do MPRDI para a imagem Barbara com blocos deslocados de múltiplos de 1/2 do seu tamanho.

Imagem Gold - com deslocamento de 1/2 comprimento do bloco

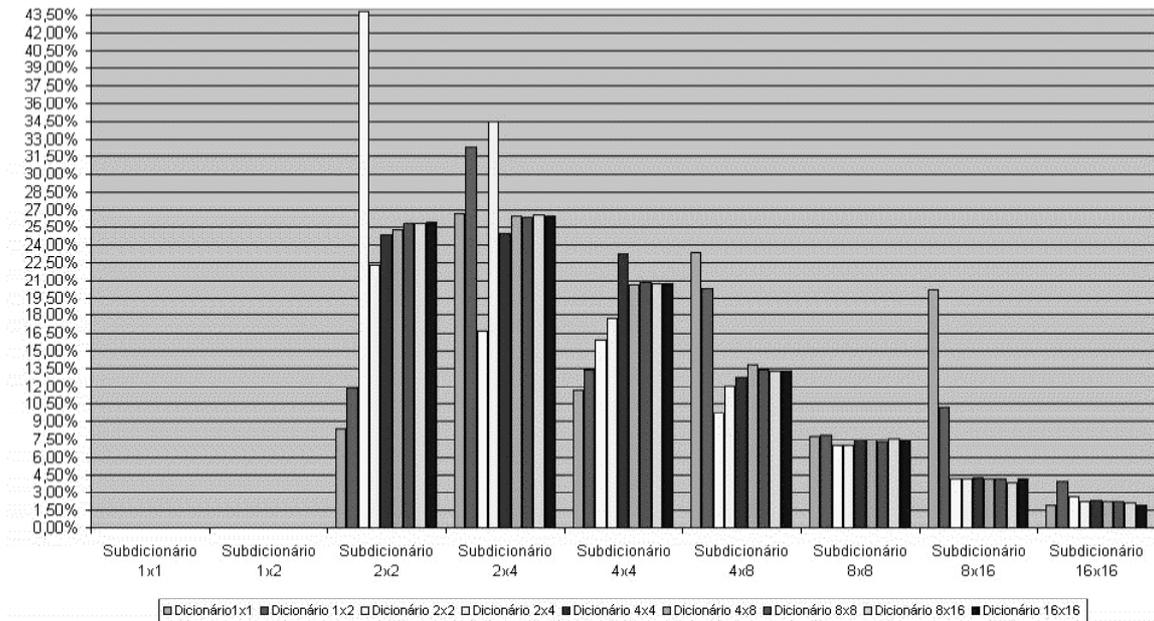


Figura 6.10: Resultado do levantamento estatístico do MPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/2 do seu tamanho.

Imagem F16 - com deslocamento de 1/2 comprimento do bloco

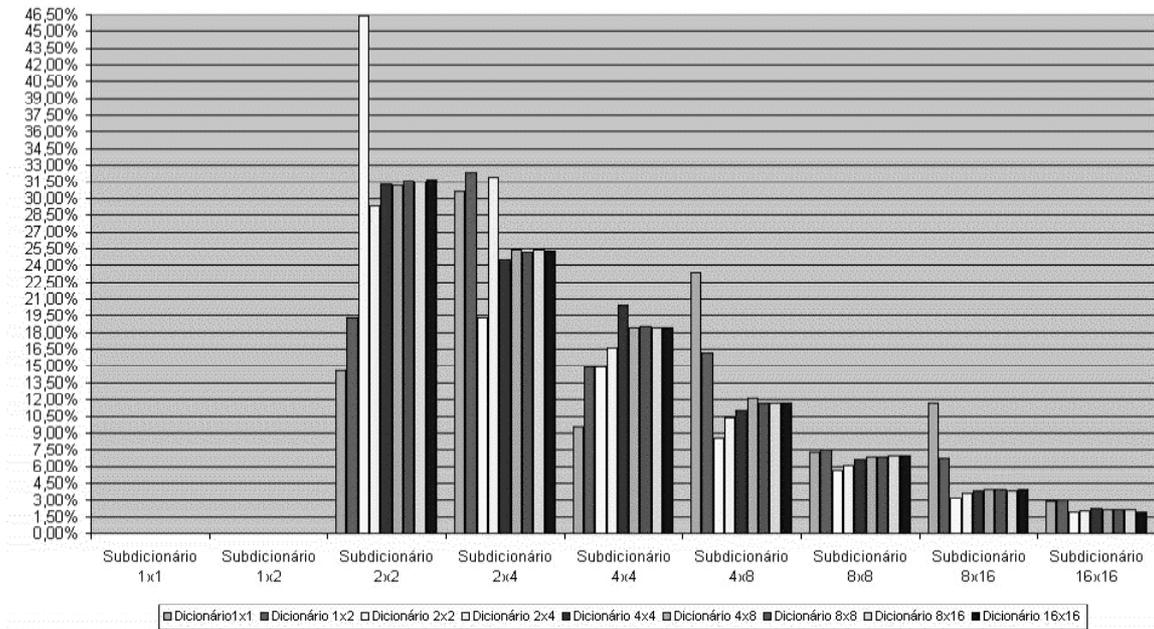


Figura 6.9: Resultado do levantamento estatístico do MMPRDI para a imagem F16 com blocos deslocados de múltiplos de 1/2 do seu tamanho.

Imagem Lena - com deslocamento de 1/2 comprimento do bloco

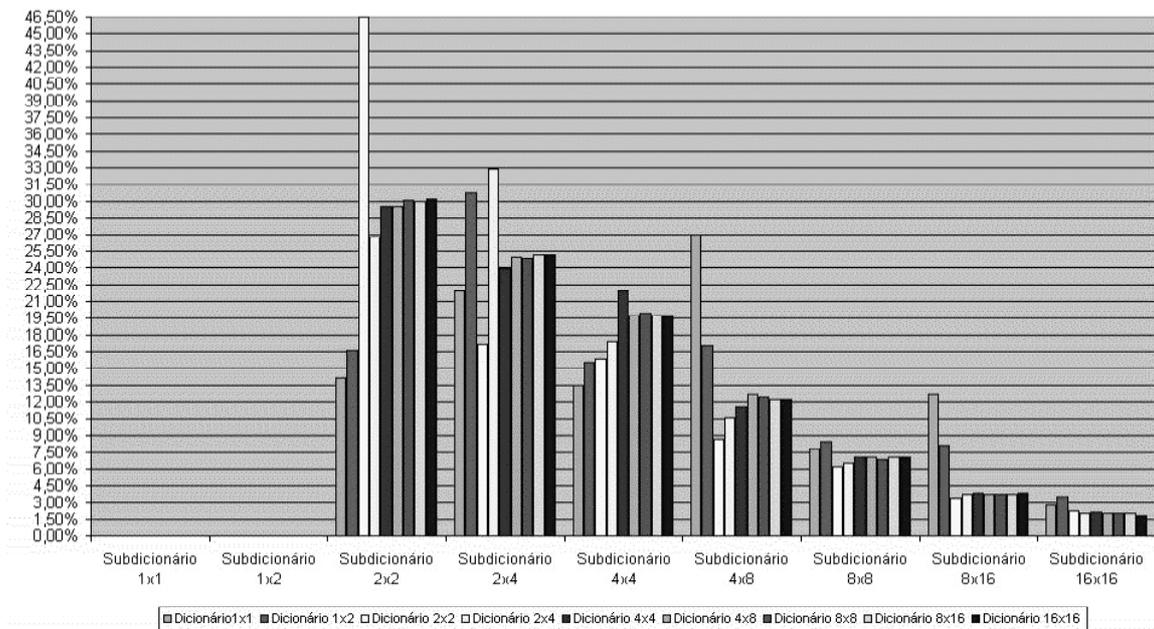


Figura 6.11: Resultado do levantamento estatístico do MMPRDI para a imagem Lena com blocos deslocados de múltiplos de 1/2 do seu tamanho.

Imagem PP1205 - com deslocamento de 1/2 comprimento do bloco

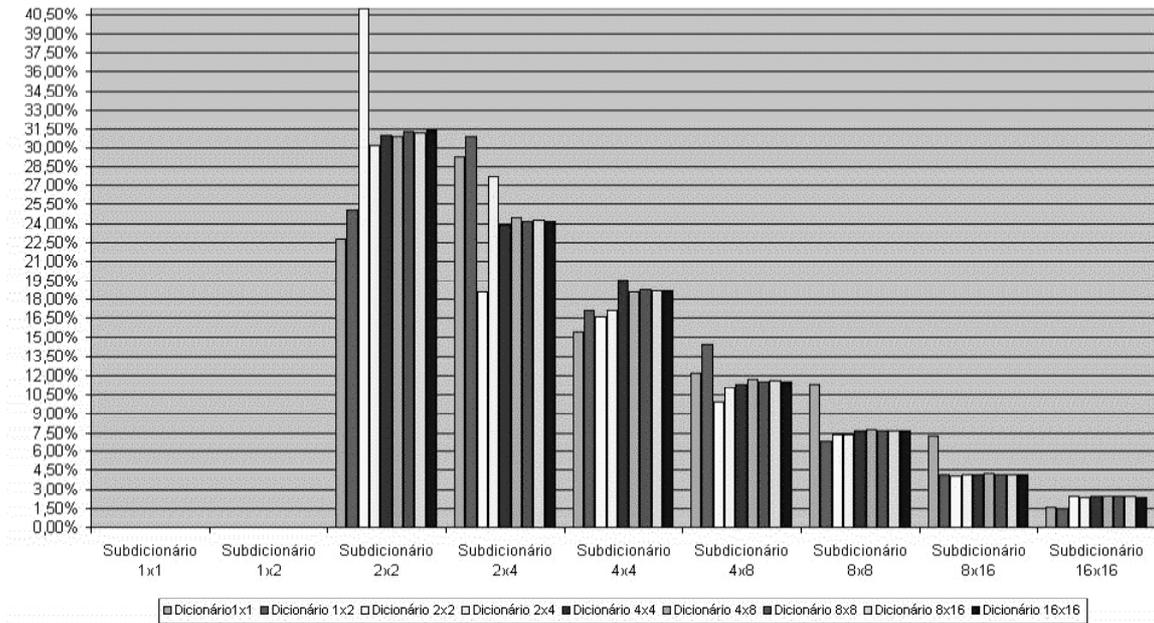


Figura 6.12: Resultado do levantamento estatístico do MPPRDI para a imagem PP1205 com blocos deslocados de múltiplos de 1/2 do seu tamanho.

Imagem PP1209 - com deslocamento de 1/2 comprimento do bloco

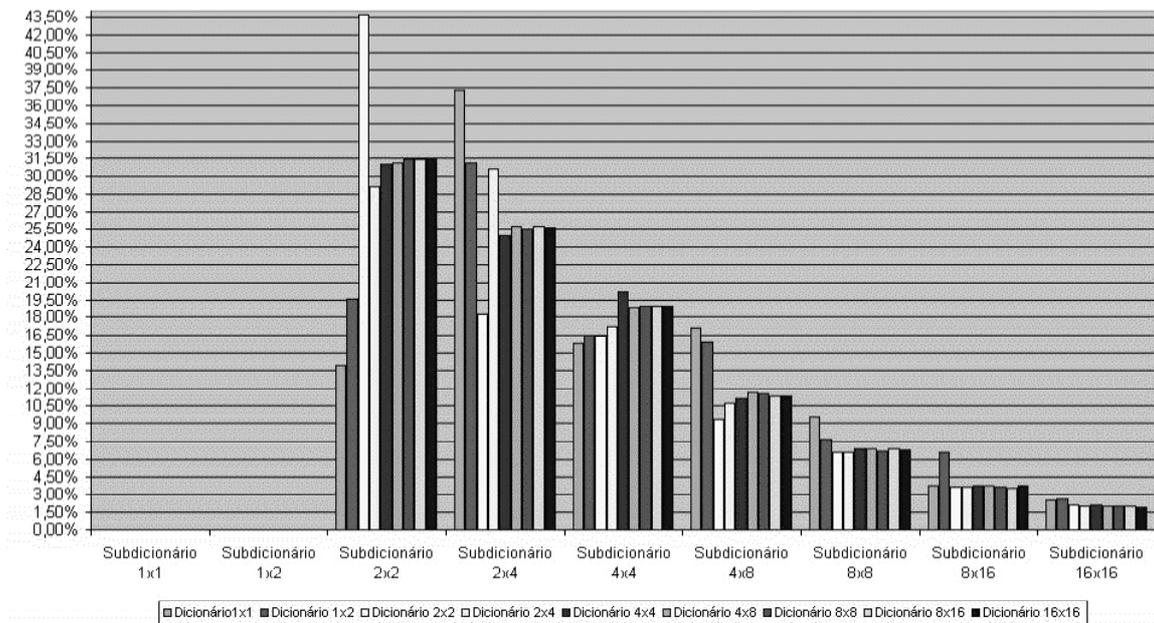


Figura 6.13: Resultado do levantamento estatístico do MPPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/2 do seu tamanho.

6.4.2 Resultados de blocos deslocados mais usados com deslocamento de múltiplos de $1/2$ de seu tamanho

Na seção 6.4.1 vimos que a escala 2×2 foi a mais utilizada para atualizar cada escala do dicionário, seguida de perto pela escala 2×4 . Esta seção mostra quais foram os tamanhos de blocos mais usados para atualizar o dicionário. As abcissas dos gráficos das figuras 6.14 a 6.19 indicam qual é a escala do dicionário que está sendo analisada, bem como o subdicionário ao qual pertence (ver capítulo 4); as ordenadas contém as probabilidades dos tamanhos de blocos que são deslocados sobre a imagem reconstruída. Ele também informa quais foram os blocos mais usados para atualizar o dicionário; as legendas informam os tamanhos de blocos usados no deslocamento sobre a imagem reconstruída (note que o bloco sem deslocamento é na verdade o bloco concatenado ver capítulo 4, os outros termos são os tamanhos dos blocos). Por exemplo, o termo “Bloco 16×16 ”. Designa um bloco com dimensão 16×16 que está sendo deslocado de múltiplos de $1/2$ do seu tamanho sobre a imagem reconstruída. Pode-se notar por estes gráficos que os dicionários 2×2 e 2×4 são constituídos de blocos deslocados de tamanhos 2×2 e 2×4 , respectivamente. Isto quer dizer que, as escalas do dicionário 2×2 e 2×4 são atualizadas só com blocos deslocados de tamanho correspondente a elas mesmas. Então, na codificação do algoritmo poderíamos considerar um flag informando a qual subdicionário pertence o bloco concatenado e um outro informando qual o tamanho do bloco foi utilizado no deslocamento (ver capítulo 4). As outras escalas do dicionário também foram atualizadas com blocos deslocados de tamanho correspondente a elas mesmas. Porém, em algumas delas, o bloco sem deslocamento (o próprio bloco concatenado usado para atualizar o dicionário) também contribuiu para a atualização de cada uma delas, isto ocorre principalmente nas escalas 16×16 , 8×16 e também nas escalas 8×8 e 4×8 . As próximas seções mostram os resultados com blocos deslocados de múltiplos de $1/4$ e $1/8$ de seu tamanho.

Barbara - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados

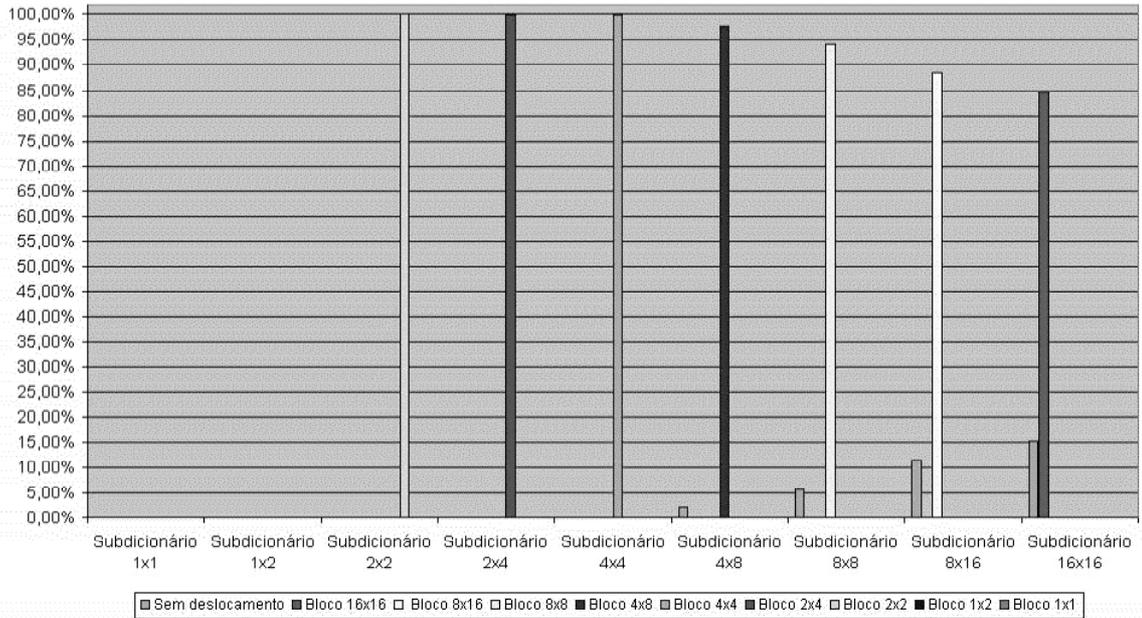


Figura 6.14: Resultado do levantamento estatístico do MPRDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.

Gold - - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados deslocado

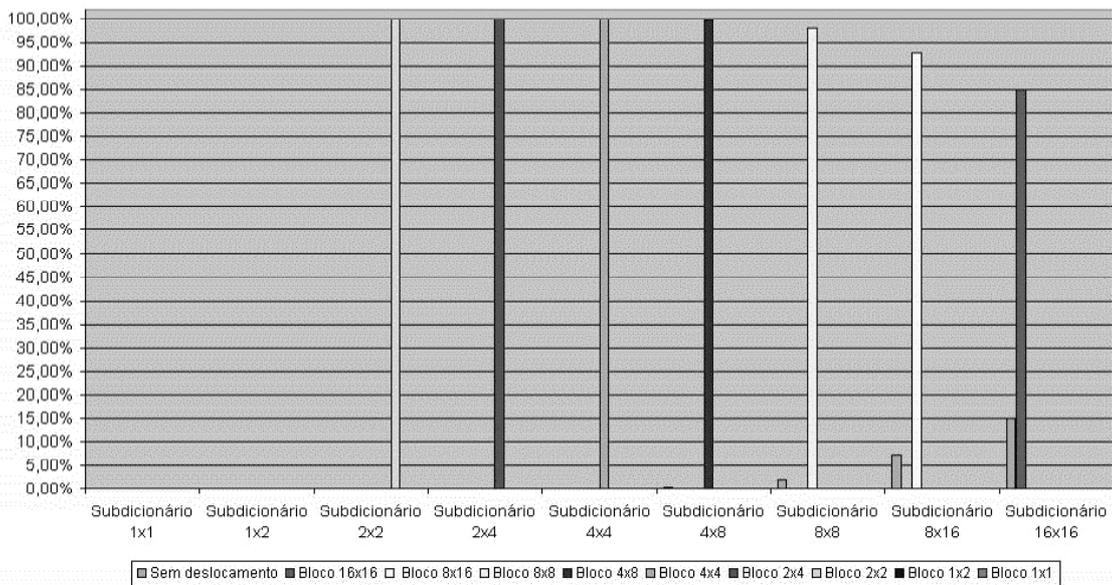


Figura 6.16: Resultado do levantamento estatístico do MPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.

F16 - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados

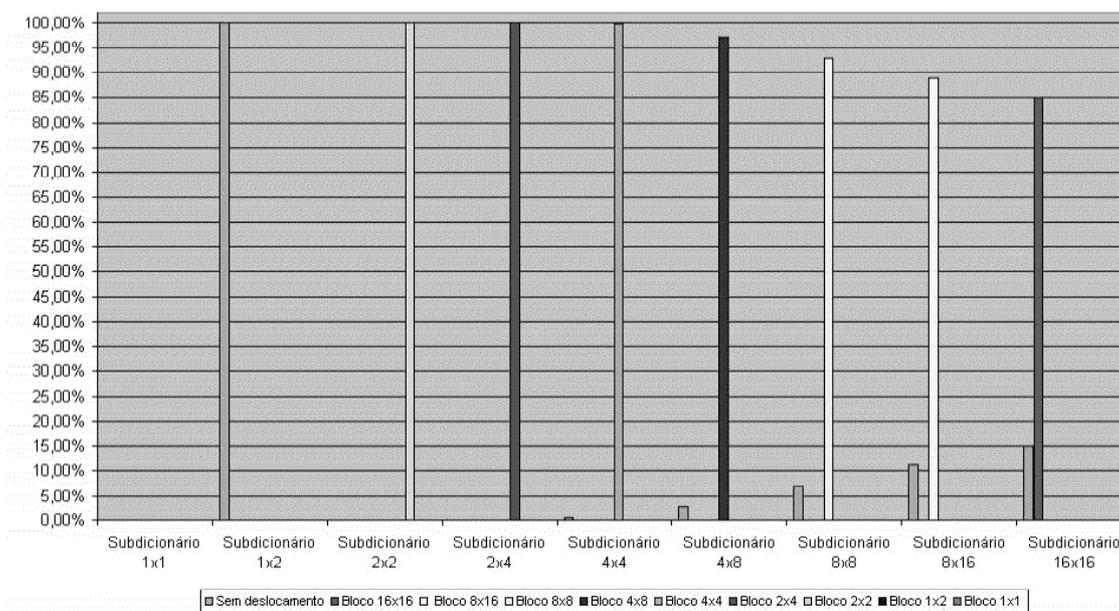


Figura 6.15: Resultado do levantamento estatístico do MPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.

Lena - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados

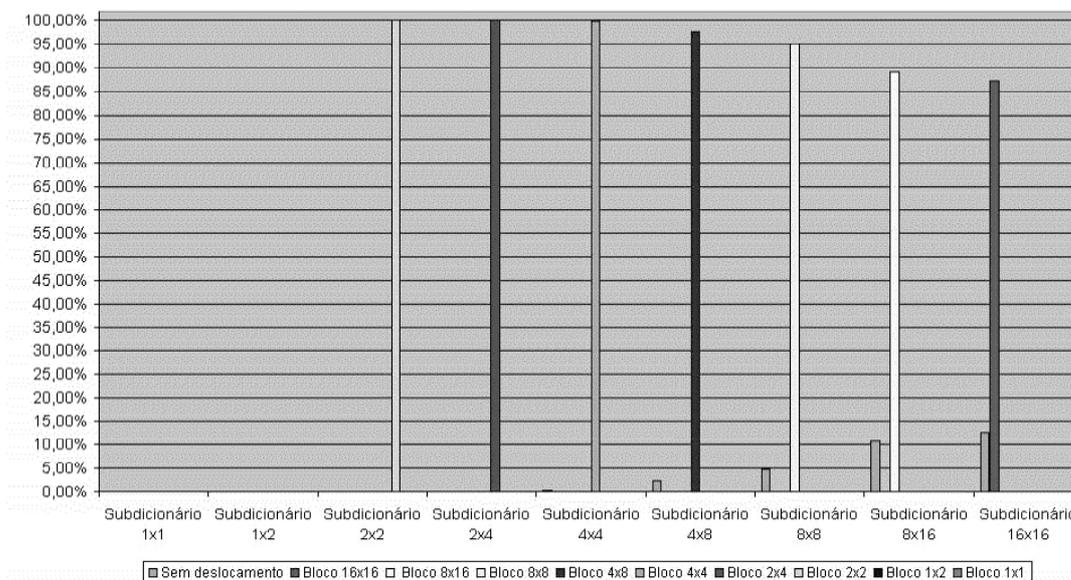


Figura 6.17: Resultado do levantamento estatístico do MPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos 1/2 do seu tamanho.

PP1205 - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados

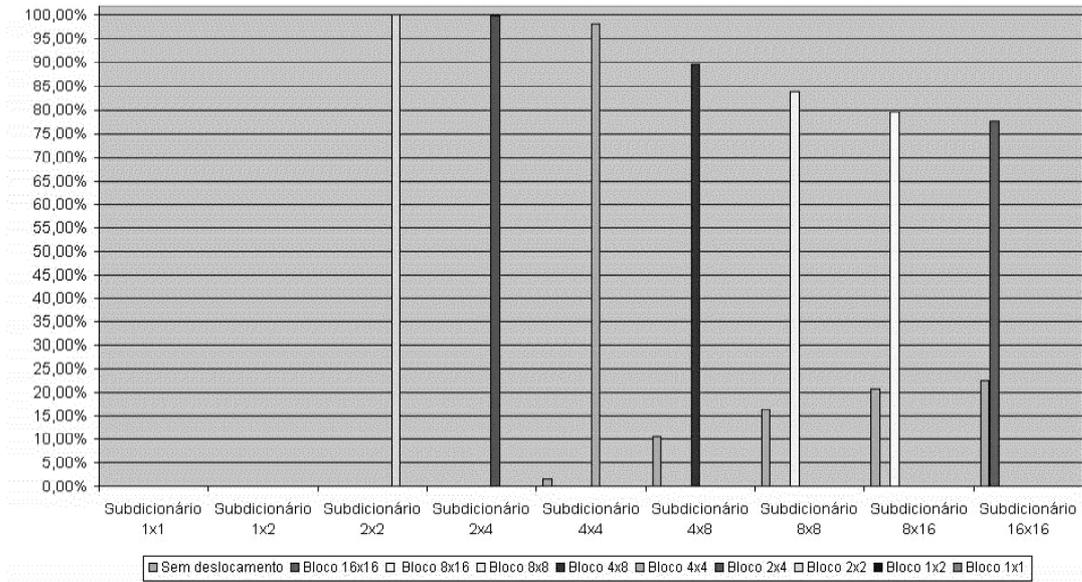


Figura 6.18: Resultado do levantamento estatístico do MPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.

PP1209 - com deslocamento de 1/2 comprimento do bloco - mostrando os blocos mais utilizados

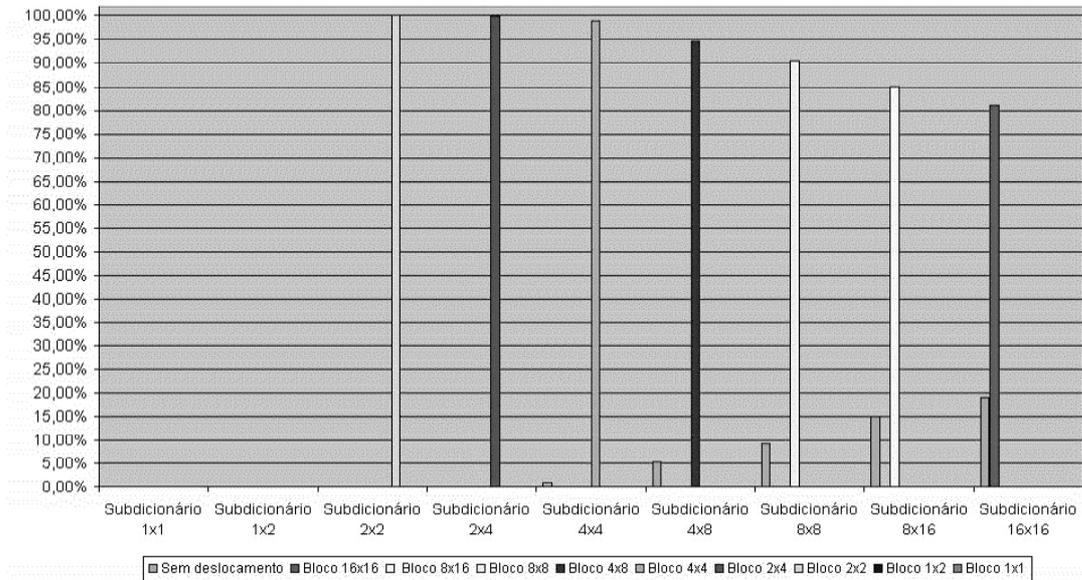


Figura 6.19: Resultado do levantamento estatístico do MPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/2 do seu tamanho.

6.4.3 Resultados para adição de blocos deslocados de múltiplos de 1/4 de seu tamanho

Vimos anteriormente o levantamento estatístico do MMPRDI com blocos deslocados de múltiplos de 1/2 de seu tamanho. Nesta seção mostramos o resultado da medição das probabilidades dos vetores do dicionário do MMPRDI incluindo deslocamento de blocos de múltiplos de 1/4 de seu tamanho. Foram utilizadas as mesmas imagens de teste da seção 6.4.1. As figuras 6.20 a 6.25 mostram os gráficos do levantamento estatístico com as probabilidades dos vetores dos subdicionários. A apresentação dos resultados segue o mesmo padrão da seção 6.4.1. Os gráficos mostram que o dicionário 4×4 foi o que mais contribuiu para a atualização de cada escala do dicionário. A seção seguinte mostra quais foram as dimensões de blocos mais usadas para as mesmas imagens utilizadas nas simulações anteriores.

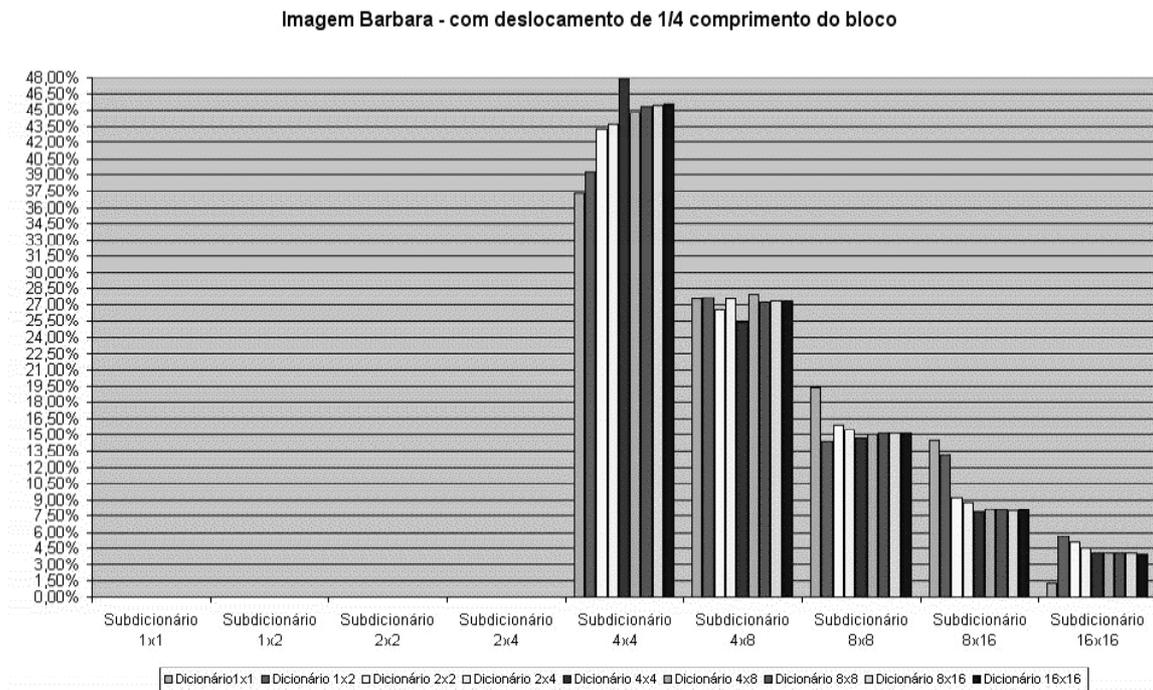


Figura 6.20: Resultado do levantamento estatístico do MMPRDI para a imagem Barbara com blocos deslocados de múltiplo de 1/4 do seu tamanho.

Imagem F16 com deslocamento de 1/4 comprimento do bloco

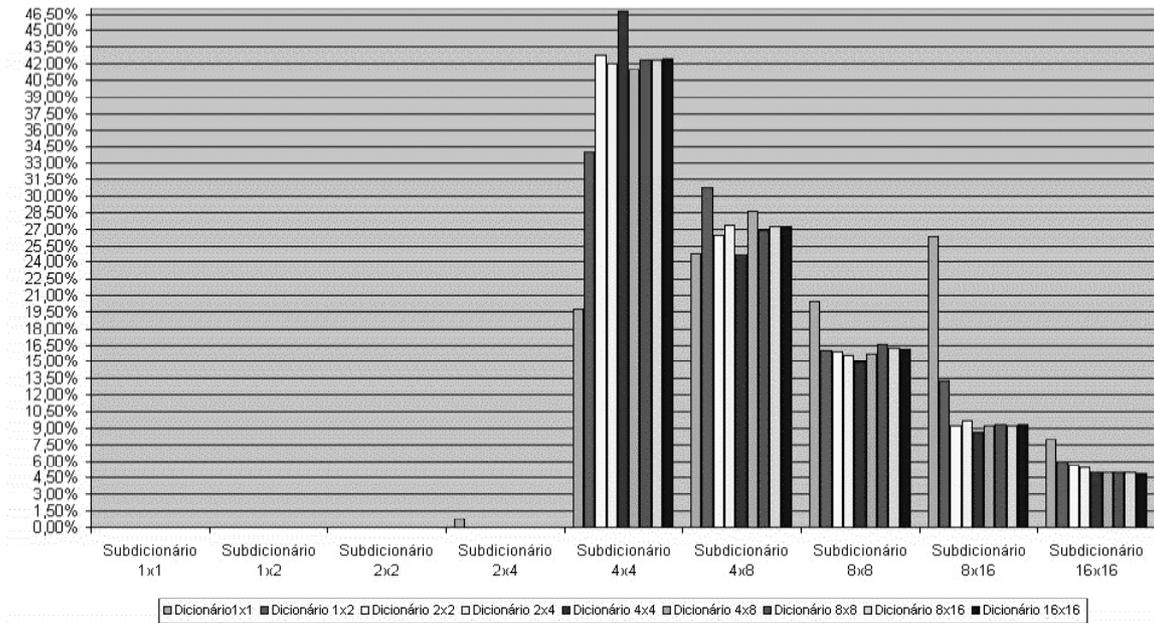


Figura 6.21: Resultado do levantamento estatístico do MPRDI para a imagem F16 com blocos deslocados de múltiplo de 1/4 do seu tamanho.

Imagem Gold com deslocamento de 1/4 comprimento do bloco

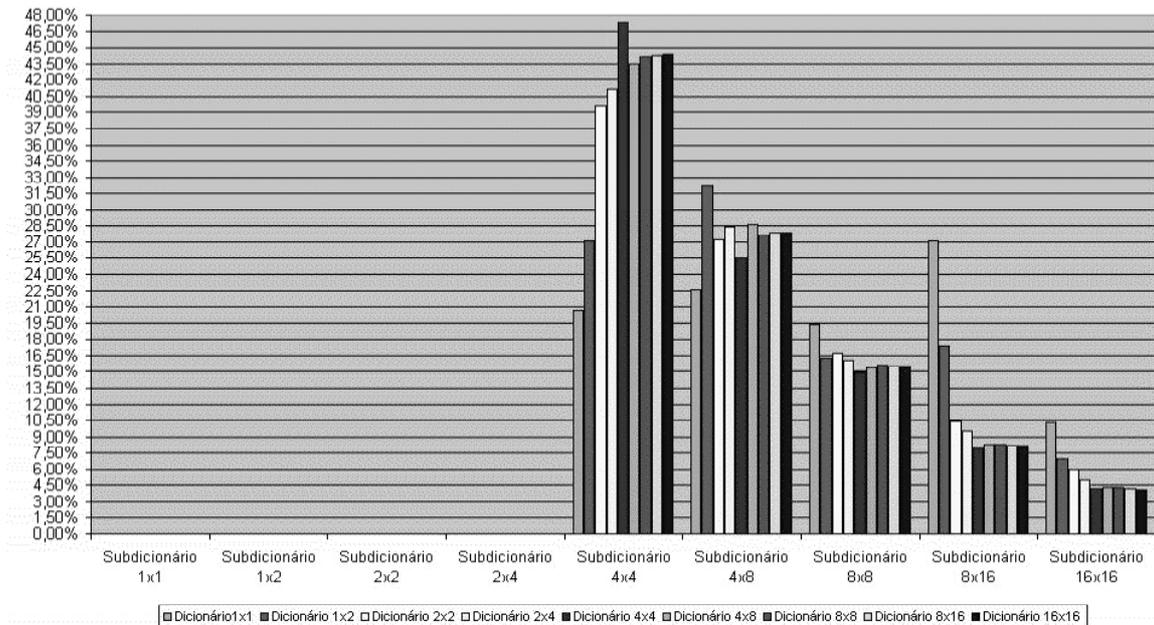


Figura 6.22: Resultado do levantamento estatístico do MPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/4 do seu tamanho.

Imagem Lena com deslocamento de 1/4 comprimento do bloco

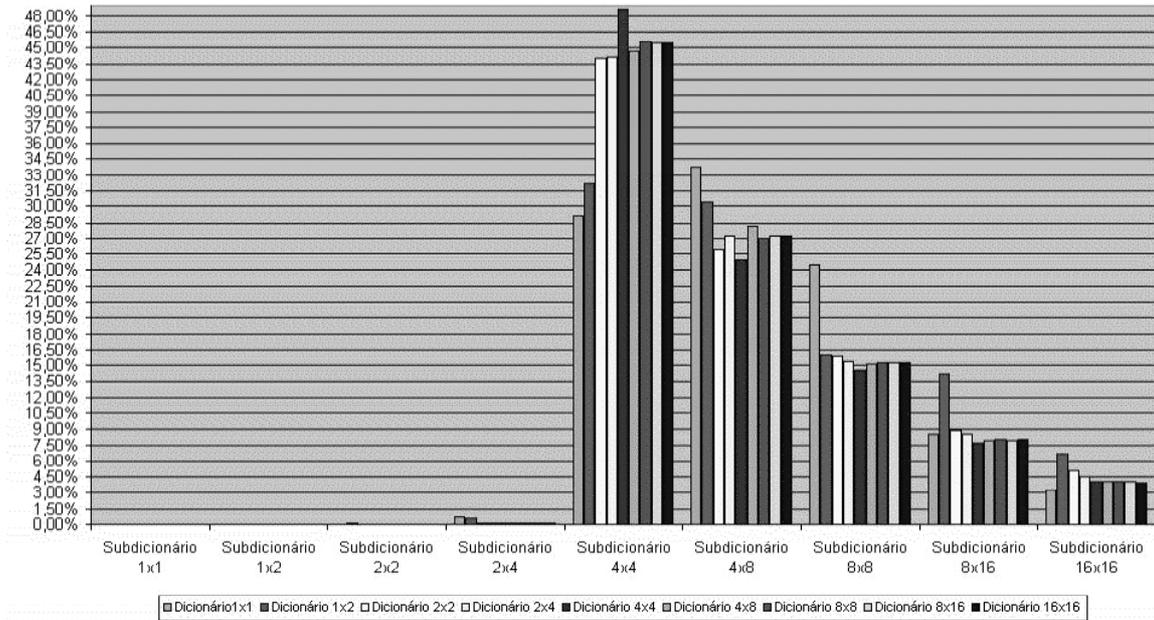


Figura 6.23: Resultado do levantamento estatístico do MPPRDI para a imagem Lena com blocos deslocados de múltiplo 1/4 do seu tamanho.

Imagem PP1205 com deslocamento de 1/4 comprimento do bloco

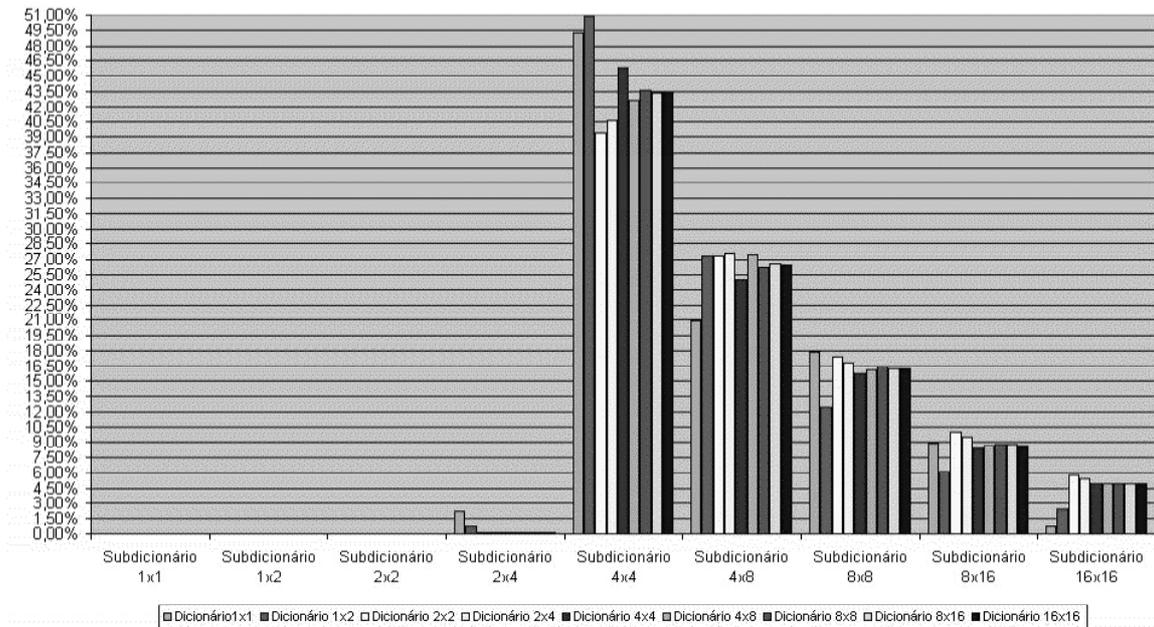


Figura 6.24: Resultado do levantamento estatístico do MPPRDI para a imagem PP1205 com blocos deslocados de múltiplos 1/4 do seu tamanho.

Imagem PP1209 com deslocamento de 1/4 comprimento do bloco

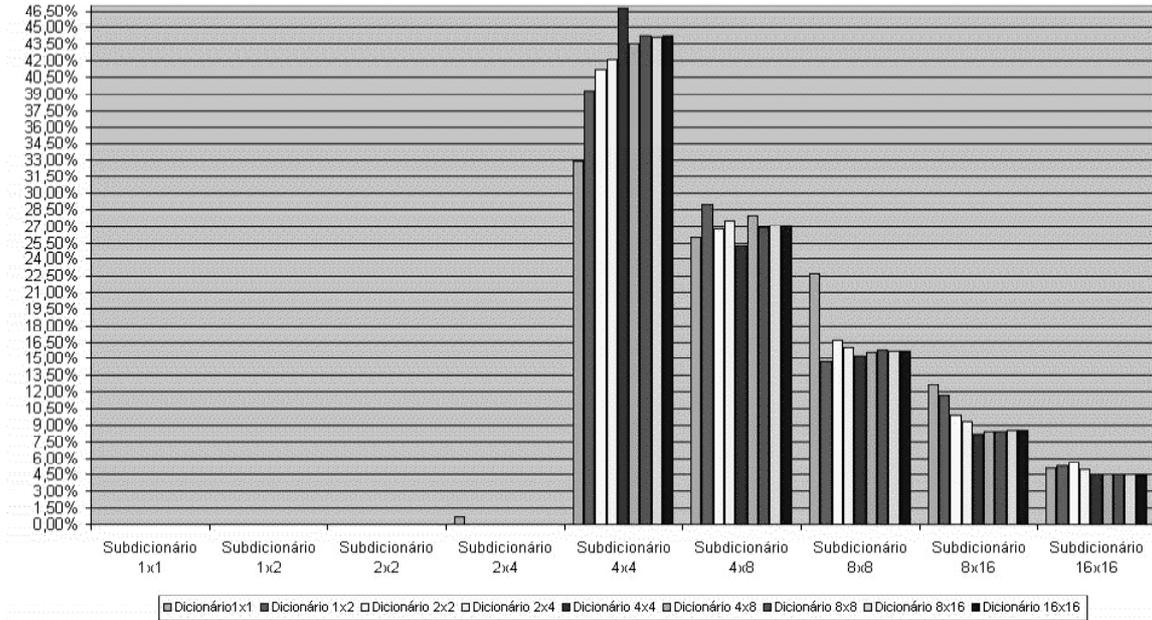


Figura 6.25: Resultado do levantamento estatístico do MPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/4 do seu tamanho.

6.4.4 Resultados de blocos deslocados mais usados com deslocamento de múltiplos de 1/4 de seu tamanho

Vimos anteriormente (seção 6.4.3), que a escala 4×4 foi a mais utilizada para atualizar cada escala do dicionário. Esta seção mostra quais foram os tamanhos de blocos mais usados para atualizar o dicionário para blocos deslocados de múltiplos de 1/4 de seu tamanho. As mesmas imagens utilizadas para teste nas seções anteriores foram também usadas nesta seção. A apresentação dos resultados segue o mesmo padrão da seção 6.4.2. Pode-se notar pelos gráficos das figuras 6.26 a 6.31 que, em alguns deles, os dicionários 1×2 , 2×2 , além de 2×4 e 4×4 são constituídos de blocos deslocados de tamanhos 1×2 , 2×2 , além do 2×4 e 4×4 , respectivamente. Isto quer dizer que, as escalas do dicionário são atualizadas só com blocos deslocados de tamanho correspondente a elas mesmas. Porém, em alguns gráficos, como é o caso das figuras 6.29 e 6.30, o dicionário 4×4 também foi constituído de bloco sem deslocamento (o próprio bloco concatenado usado para atualizar o dicionário, ver seção 6.4.2). Então, na codificação do algoritmo poderíamos considerar

um flag informando a qual subdicionário pertence o bloco concatenado e um outro informando qual o tamanho do bloco foi utilizado no deslocamento (ver capítulo 4). As outras escalas do dicionário também foram atualizadas com blocos deslocados de tamanho correspondente a elas mesmas. Porém, em algumas delas, o bloco sem deslocamento também contribuiu para a atualização de cada uma delas, isto ocorre principalmente nas escalas 16×16 , 8×16 e também nas escalas 8×8 e 4×8 . As próximas seções mostram os resultados com blocos deslocados de múltiplos de $1/8$ de seu tamanho.

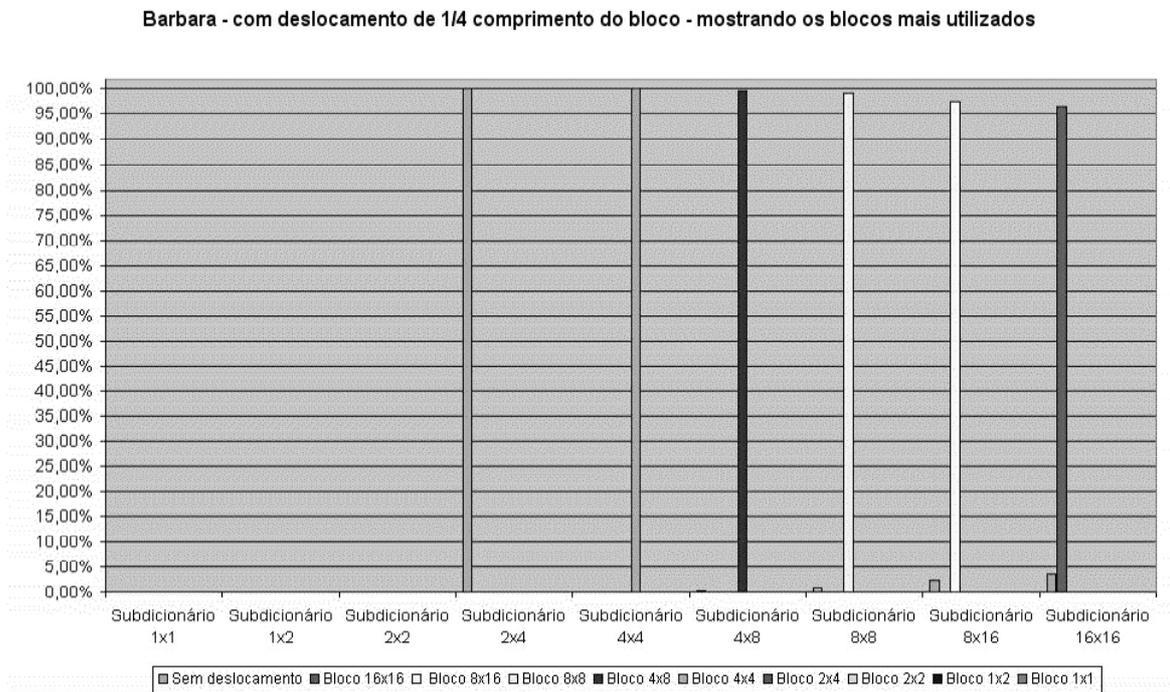


Figura 6.26: Resultado do levantamento estatístico do MMPDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos de $1/4$ do seu tamanho.

F16 - com deslocamento de 1/4 comprimento do bloco - mostrando os blocos mais utilizados

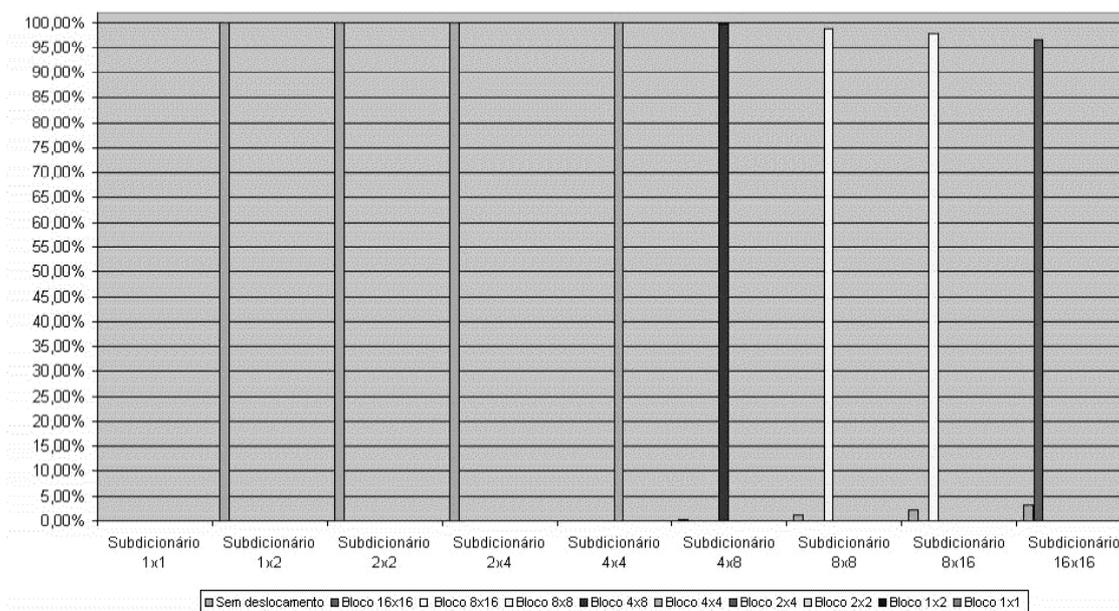


Figura 6.27: Resultado do levantamento estatístico do MPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.

Gold - com deslocamento de 1/4 comprimento do bloco - mostrando os blocos mais utilizados

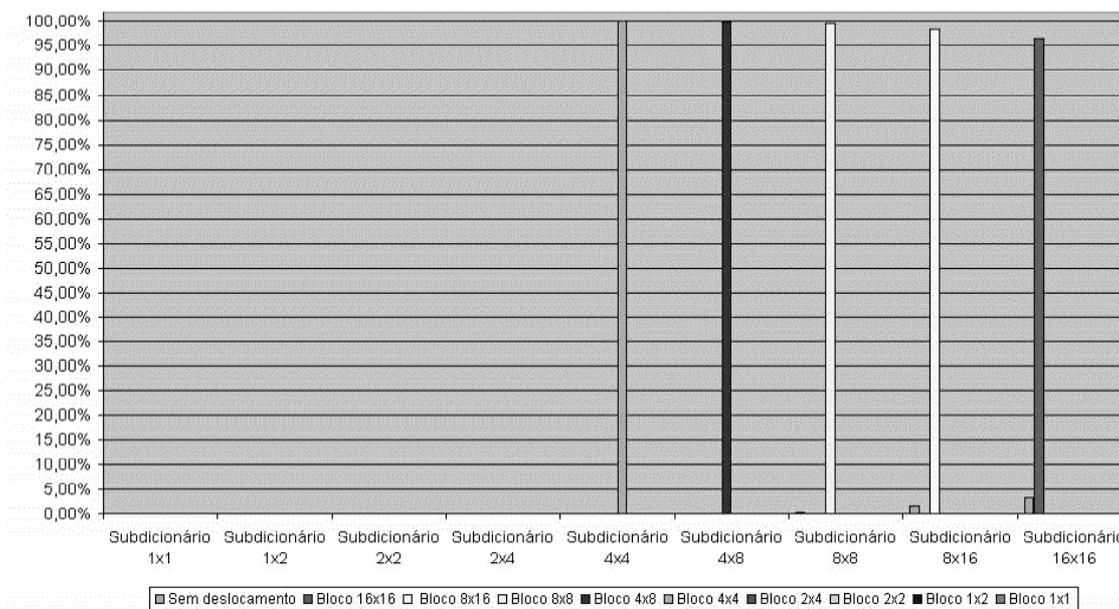


Figura 6.28: Resultado do levantamento estatístico do MPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.

Lena - com deslocamento de 1/4 comprimento do bloco - mostrando os blocos mais utilizados

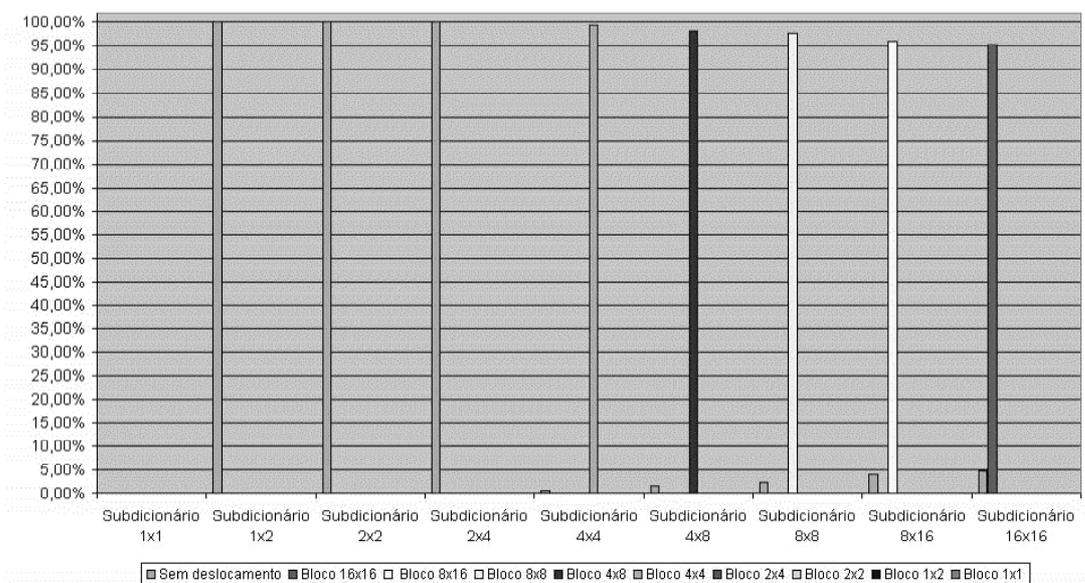


Figura 6.29: Resultado do levantamento estatístico do MPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.

PP1205 - com deslocamento de 1/4 comprimento do bloco - mostrando os blocos mais utilizados

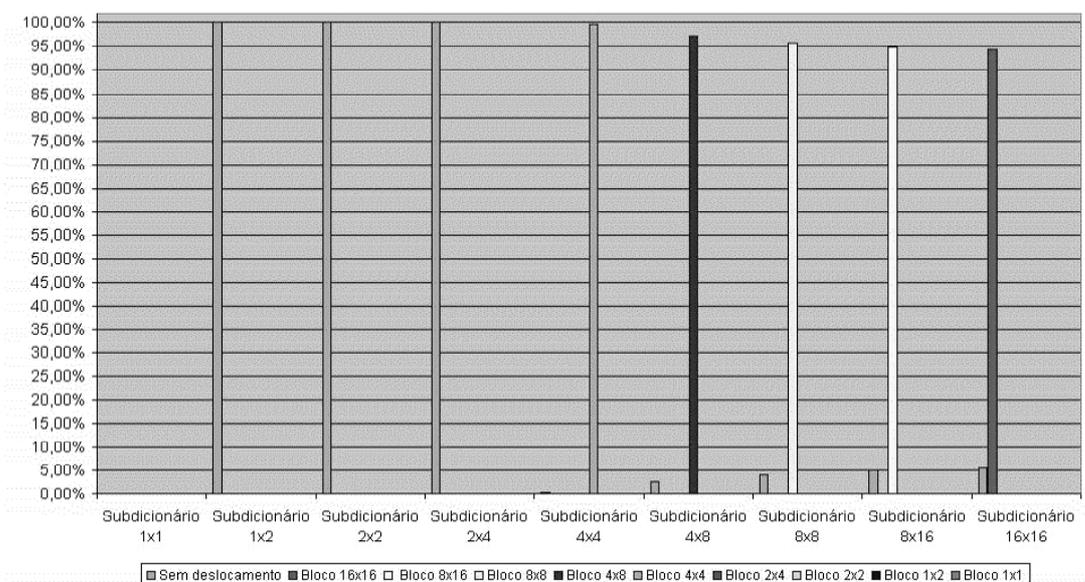


Figura 6.30: Resultado do levantamento estatístico do MPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.

PP1209 - com deslocamento de 1/4 comprimento do bloco - mostrando os blocos mais utilizados

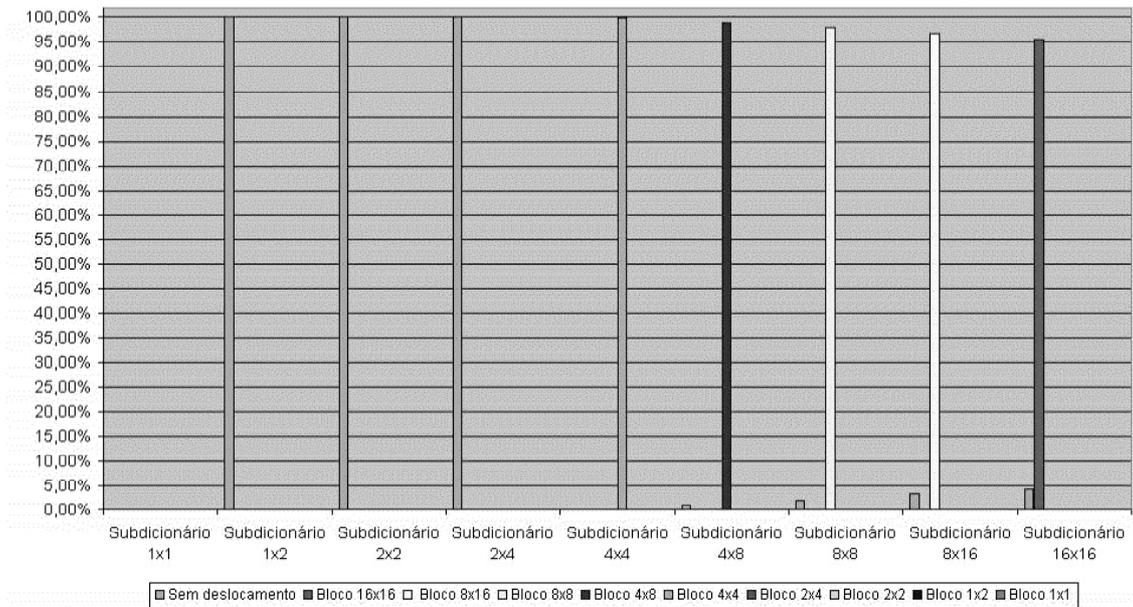


Figura 6.31: Resultado do levantamento estatístico do MPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/4 do seu tamanho.

6.4.5 Resultados para adição de blocos deslocados de múltiplos de 1/8 de seu tamanho

Vimos anteriormente o levantamento estatístico do MPRDI com blocos deslocados de múltiplos de 1/8 de seu tamanho. Nesta seção mostramos o resultado da medição das probabilidades dos vetores do dicionário do MPRDI incluindo deslocamento de blocos de múltiplos de 1/8 de seu tamanho. Foram utilizadas as mesmas imagens de teste da seção 6.4.1. As figuras 6.32 a 6.37 mostram os gráficos do levantamento estatístico com as probabilidades dos vetores dos subdicionários. A apresentação dos resultados segue o mesmo padrão da seção 6.4.1. Os gráficos mostram que o dicionário 4×4 foi o que mais contribuiu para a atualização de cada escala do dicionário. A seção seguinte mostra quais foram as dimensões de blocos mais usadas para as mesmas imagens utilizadas nas simulações.

Imagem Barbara - com deslocamento de 1/8 comprimento do bloco

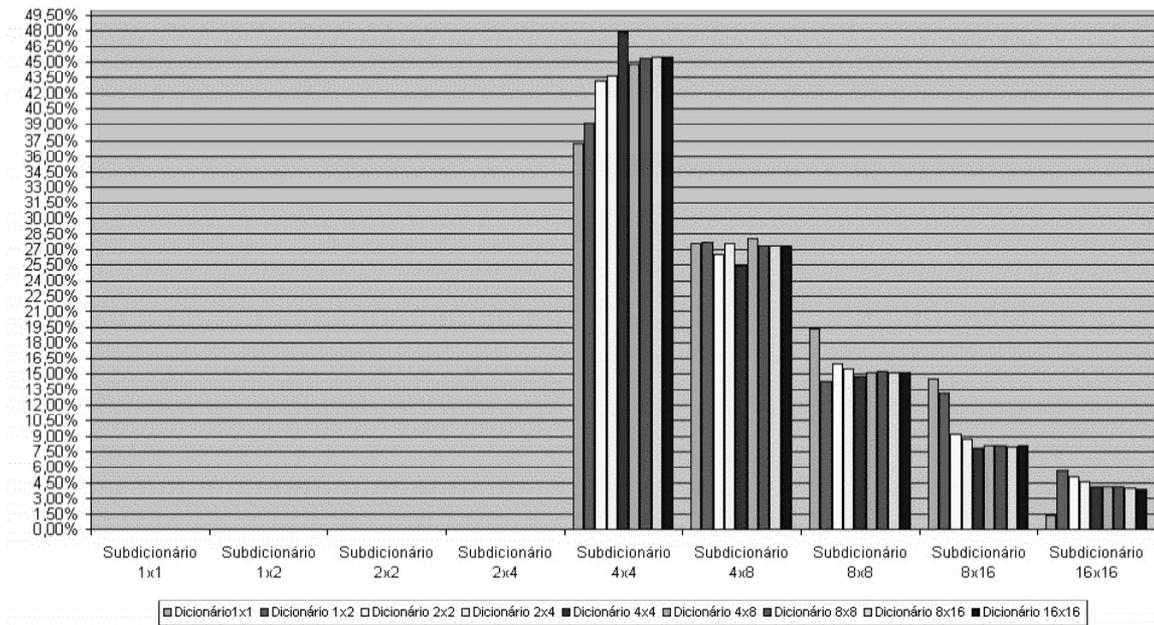


Figura 6.32: Resultado do levantamento estatístico do MPPRDI para a imagem Barbara com blocos deslocados de múltiplos de 1/8 do seu tamanho.

Imagem Gold - com deslocamento de 1/8 comprimento do bloco

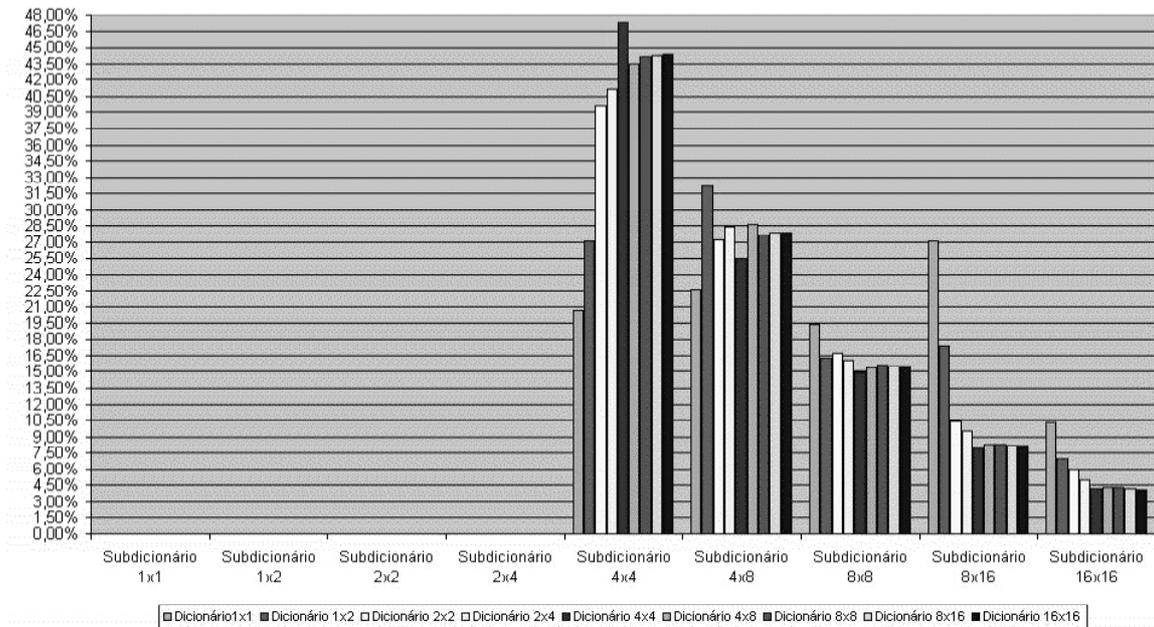


Figura 6.34: Resultado do levantamento estatístico do MPPRDI para a imagem Gold com blocos deslocados de múltiplos de 1/8 do seu tamanho.

Imagem F16 - com deslocamento de 1/8 comprimento do bloco

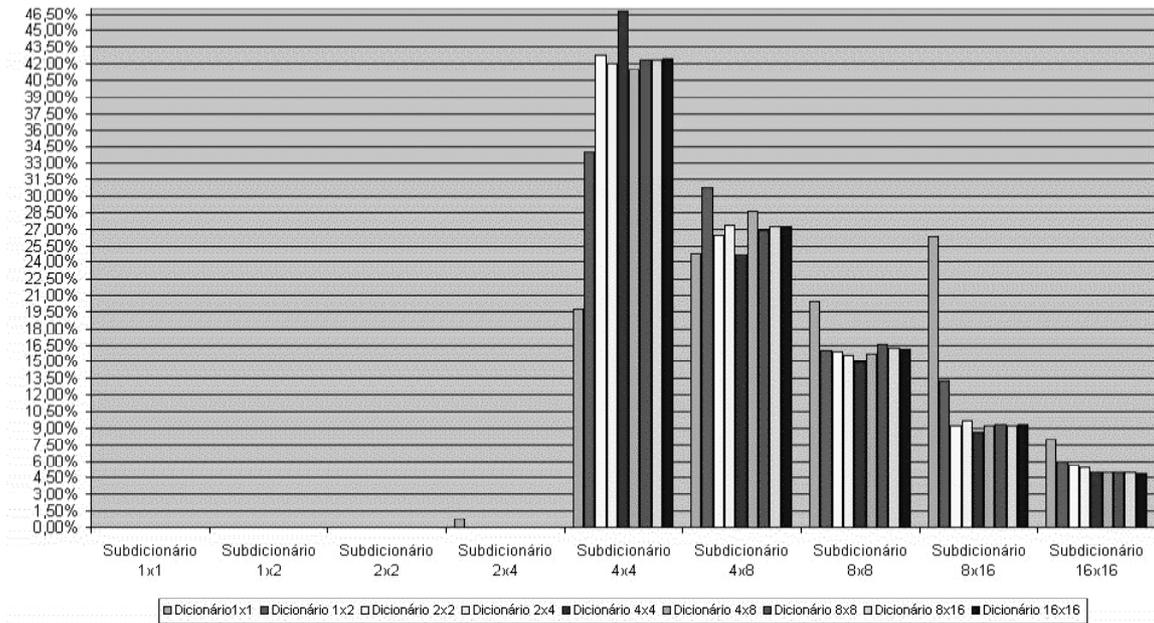


Figura 6.33: Resultado do levantamento estatístico do MPRDI para a imagem F16 com blocos deslocados de múltiplos de 1/8 do seu tamanho.

Imagem Lena - com deslocamento de 1/8 comprimento do bloco

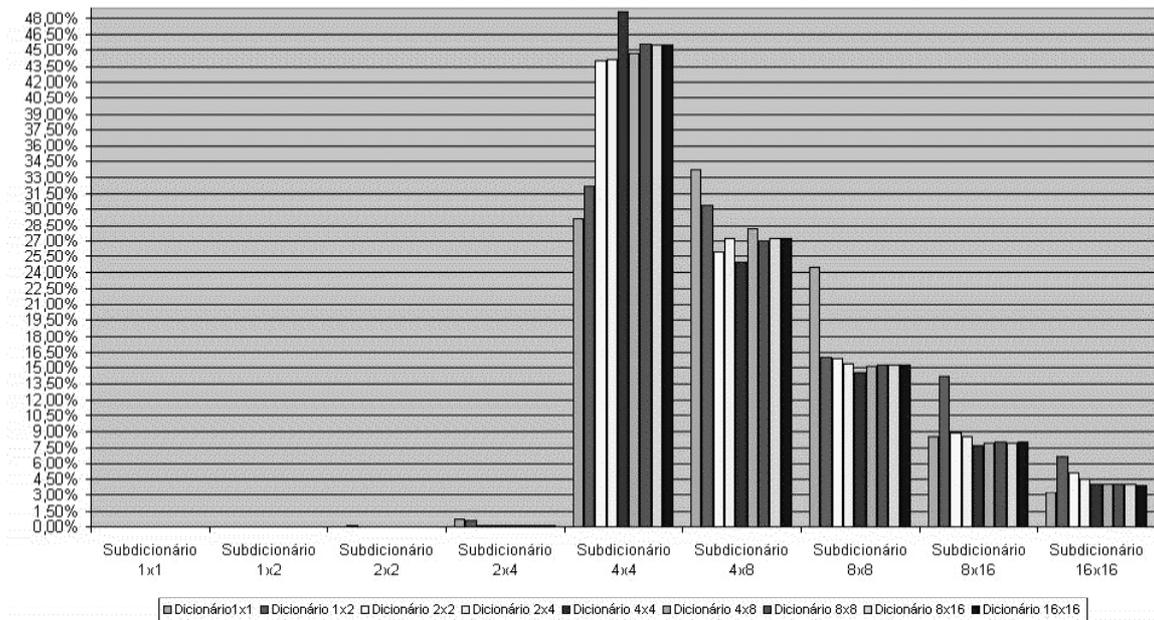


Figura 6.35: Resultado do levantamento estatístico do MPRDI para a imagem Lena com blocos deslocados de múltiplos de 1/8 do seu tamanho.

Imagem PP1205 - com deslocamento de 1/8 comprimento do bloco

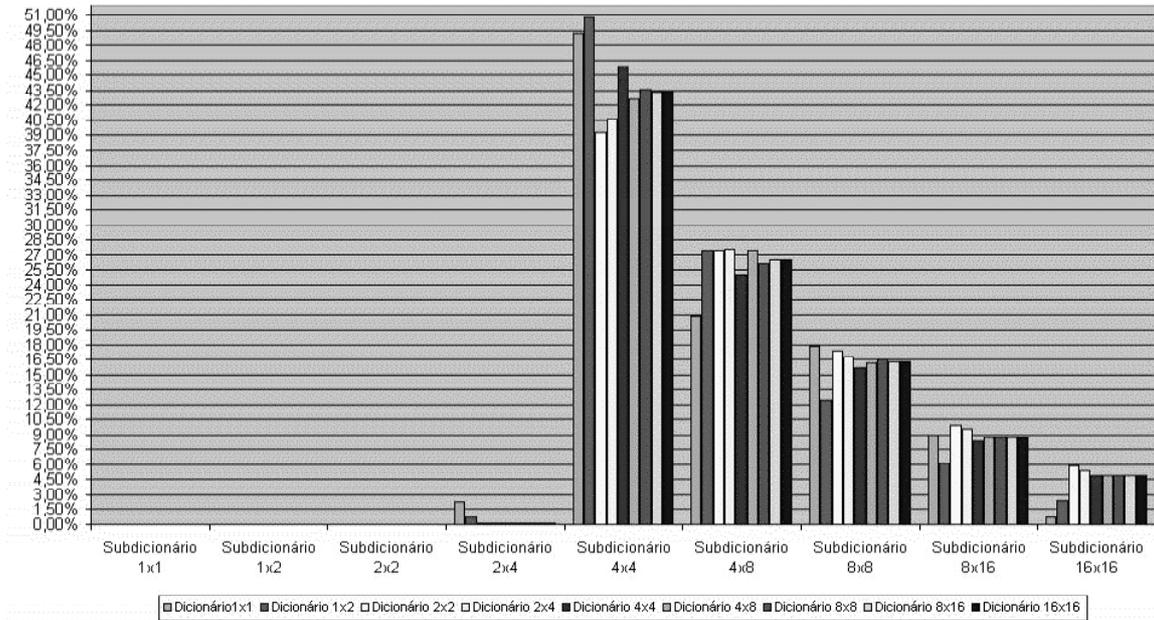


Figura 6.36: Resultado do levantamento estatístico do MPRDI para a imagem PP1205 com blocos deslocados de múltiplos de 1/8 do seu tamanho.

Imagem PP1209 - com deslocamento de 1/8 comprimento do bloco

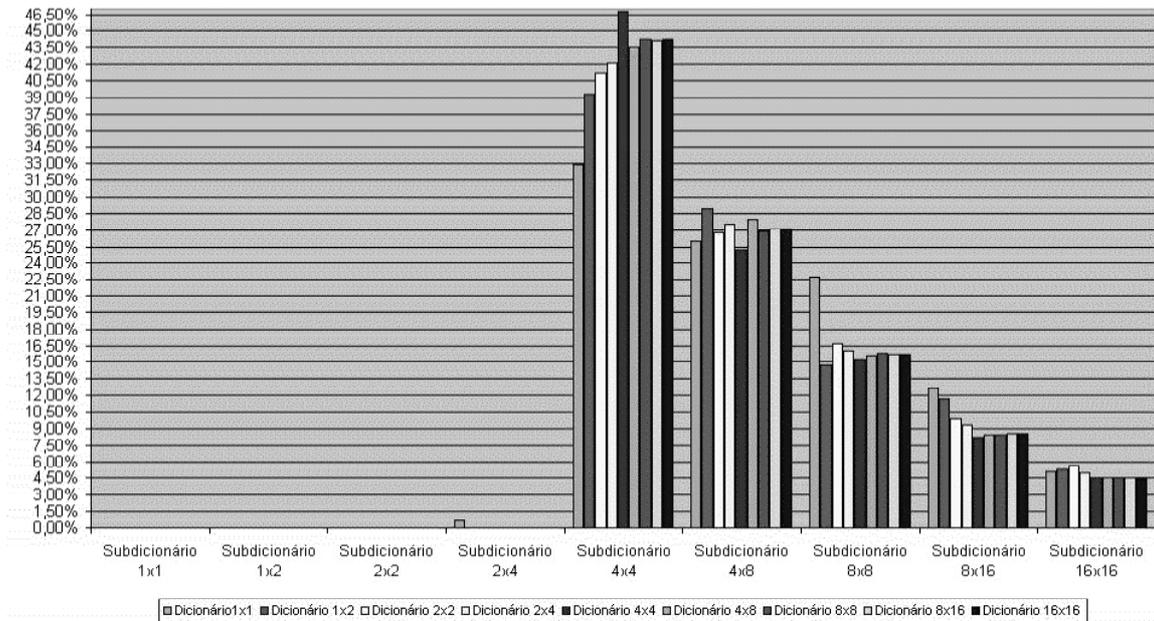


Figura 6.37: Resultado do levantamento estatístico do MPRDI para a imagem PP1209 com blocos deslocados de múltiplos de 1/8 do seu tamanho.

6.4.6 Resultados de blocos deslocados mais usados com deslocamento de múltiplos de $1/8$ de seu tamanho

Vimos anteriormente (seção 6.4.5), que a escala 4×4 foi a mais utilizada para atualizar cada escala do dicionário. Esta seção mostra quais foram os tamanhos de blocos mais usados para atualizar o dicionário para blocos deslocados de múltiplos de $1/8$ de seu tamanho. A apresentação dos resultados segue o mesmo padrão da seção 6.4.2. Pode-se notar pelos gráficos das figuras 6.26 a 6.31 que, em alguns deles, os dicionários 1×2 , 2×2 , além de 2×4 , 4×4 e também 4×8 são constituídos de blocos deslocados de tamanhos 1×2 , 2×2 , além do 2×4 , 4×4 e também 4×8 , respectivamente. Isto quer dizer que, as escalas do dicionário são atualizadas só com blocos deslocados de tamanho correspondente a elas mesmas. Então, na codificação do algoritmo poderíamos considerar um flag informando a qual subdicionário pertence o bloco concatenado e um outro informando qual o tamanho do bloco foi utilizado no deslocamento (ver capítulo 4). As outras escalas do dicionário também foram atualizadas com blocos deslocados de tamanho correspondente a elas mesmas. Porém, em algumas delas, o bloco sem deslocamento também contribuiu para a atualização de cada uma delas, isto ocorre principalmente nas escalas 16×16 , 8×16 e também nas escalas 8×8 . Na próxima seção apresentamos a conclusão dos resultados dos levantamentos estatísticos feitos para cada tipo de deslocamento de blocos.

Barbara - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

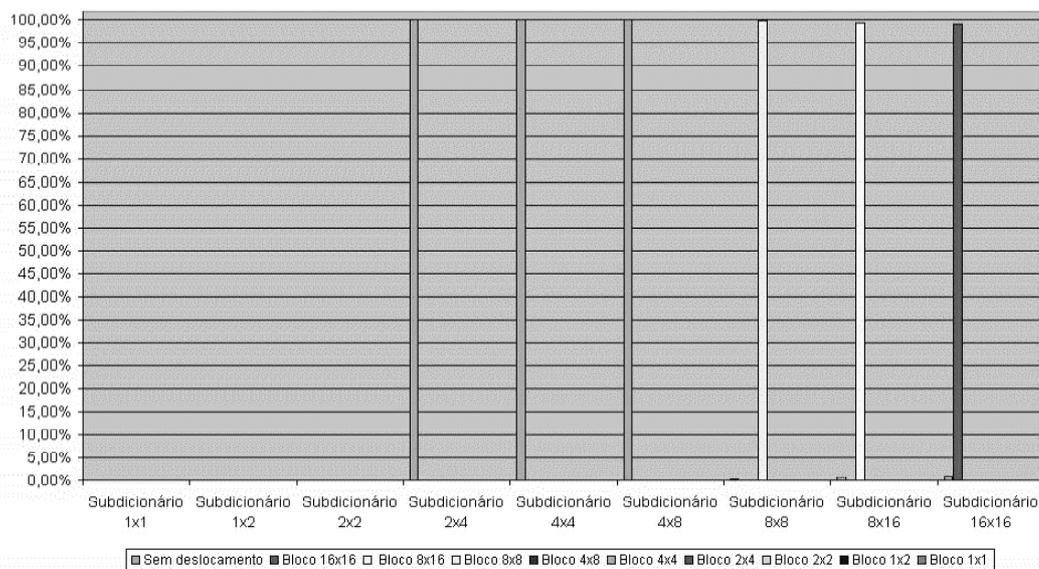


Figura 6.38: Resultado do levantamento estatístico do MPRDI para a imagem Barbara para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

F16 - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

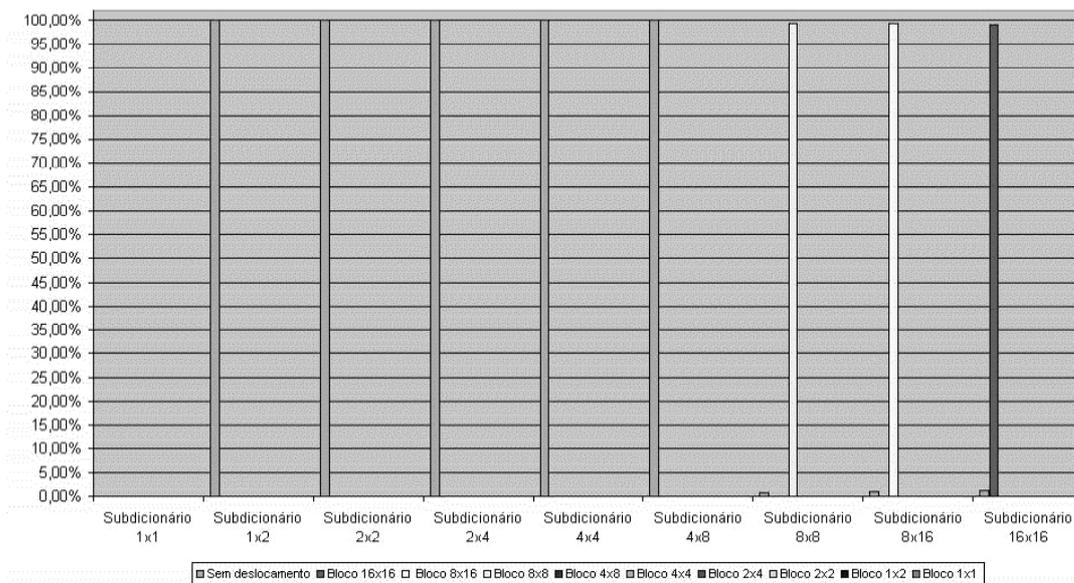


Figura 6.39: Resultado do levantamento estatístico do MPRDI para a imagem F16 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

Gold - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

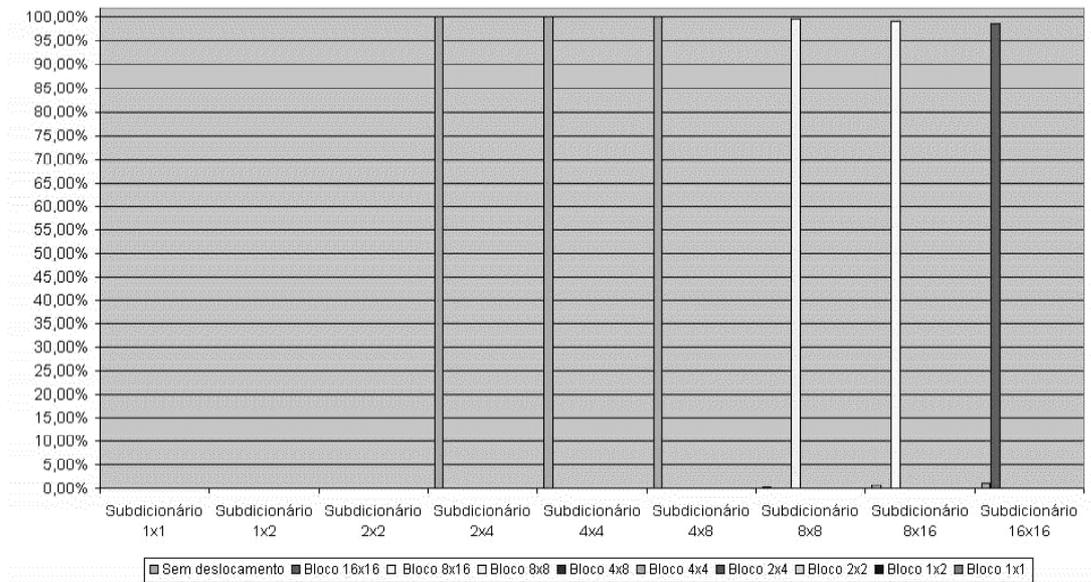


Figura 6.40: Resultado do levantamento estatístico do MPRDI para a imagem Gold para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

Lena - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

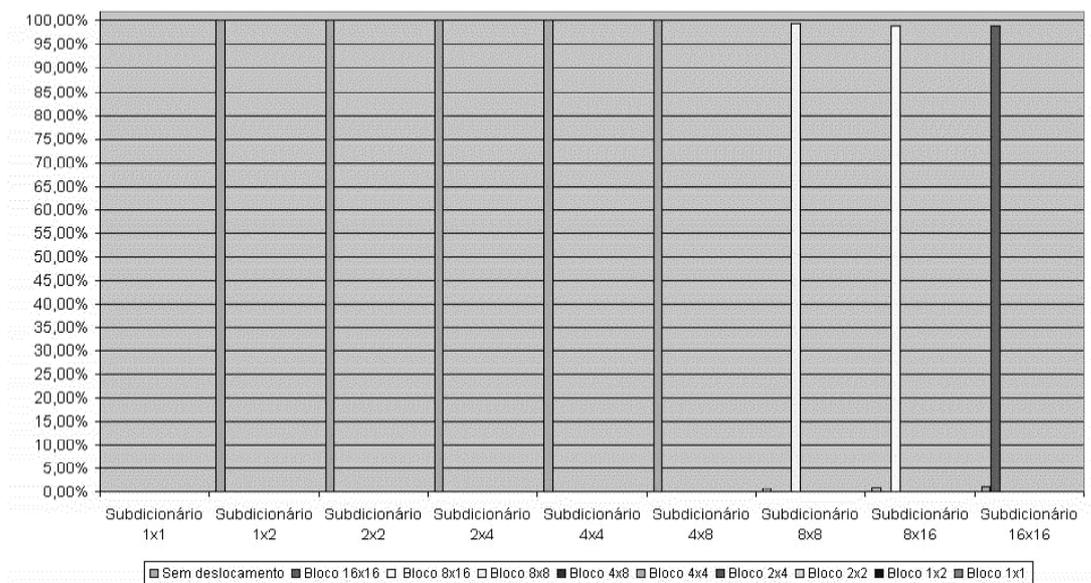


Figura 6.41: Resultado do levantamento estatístico do MPRDI para a imagem Lena para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

PP1205 - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

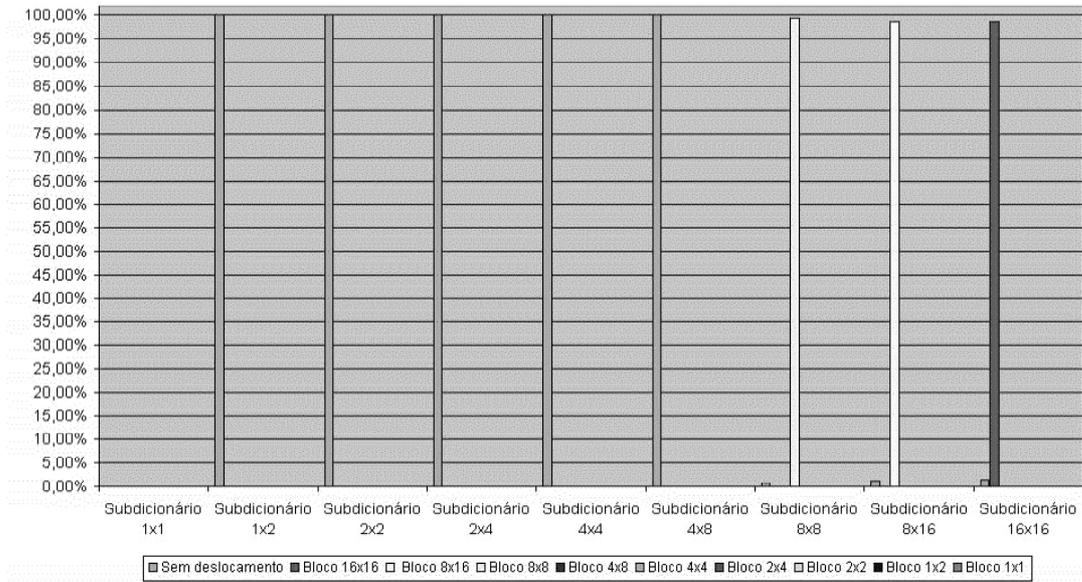


Figura 6.42: Resultado do levantamento estatístico do MPRDI para a imagem PP1205 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

PP1209 - com deslocamento de 1/8 comprimento do bloco - mostrando os blocos mais utilizados

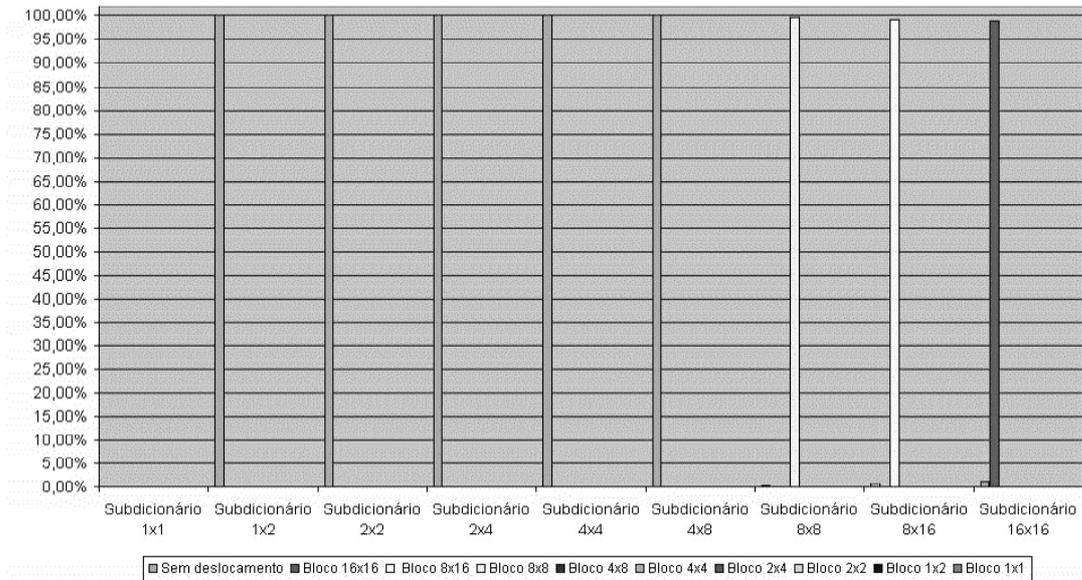


Figura 6.43: Resultado do levantamento estatístico do MPRDI para a imagem PP1209 para blocos deslocados mais usados com deslocamento de múltiplos de 1/8 do seu tamanho.

6.4.7 Levantamento Estatístico: conclusões

Pelos gráficos nas figuras 6.8 a 6.13, 6.20 a 6.25 e nas figuras 6.32 a 6.37, podemos perceber que será efetivo o uso da segmentação no espaço de código, pois as probabilidades dos vetores pertencentes ao subdicionário são bem diferentes (ver capítulo 4). O mesmo podemos observar pelos gráficos das figuras 6.14 a 6.19, 6.26 a 6.31 e das figuras 6.38 a 6.43. Assim, no primeiro caso, usaríamos na codificação um flag informando a qual subdicionário a palavra (bloco concatenado) pertence e um índice indicando sua localização dentro do subdicionário. No outro, poderíamos usar um flag informando a qual subdicionário pertence a palavra e um flag para qual tipo de deslocamento foi usado na codificação. Porém por restrição de tempo não foi possível implementar o algoritmo com uso de segmentação no espaço de códigos para deslocamentos de blocos.

Capítulo 7

Conclusão

Nesta dissertação, verificamos o desempenho para codificação de esquemas de compressão de sinais MMP, para, a partir daí, propor modificações que pudessem redundar em melhoras de desempenho. Este trabalho foi desenvolvido em três etapas.

A primeira etapa foi realizada no capítulo 4 onde avaliamos o MMP verificando a frequência de utilização das escalas do dicionário na codificação e decodificação de imagens suaves e mistas. Verificou-se que, dependendo da escala utilizada, os vetores têm diferentes probabilidades de terem sido incluídos no dicionário vindo como expansão ou contração de determinadas escalas. Estes resultados mostram que vale a pena codificar uma palavra do dicionário segmentando cada dicionário em subdicionários, cada subdicionário contendo apenas as palavras que entraram no dicionário vindas de uma determinada escala. Em outras palavras, ao invés de se codificar os índices do dicionário diretamente, se codifica um flag indicando a qual subdicionário ela pertence, e depois um índice a localizando dentro do subdicionário. Os resultados taxa-distorção deste tipo de segmentação do dicionário no espaço de códigos mostraram uma melhora significativa de desempenho (0.5dB para LENA a 0.5bpp).

A segunda etapa é descrita no capítulo 5, onde incluímos no dicionário blocos rotacionados e transpostos. Se verificou uma discreta melhora na relação sinal ruído em algumas imagens e uma piora em outras, o que indica que o crescimento do dicionário desta forma não redundava em adição de padrões que sejam relevantes. Foi feita também a avaliação se a segmentação no espaço de códigos valeria a pena para

estes dicionários, através de um levantamento estatístico dos blocos/escalas mais usados. Se concluiu que o uso deste tipo de segmentação do dicionário não seria vantajoso quando considerarmos na codificação qual o tipo de rotação ou transposta que foi utilizado. Porém, se considerarmos a segmentação utilizando o mesmo raciocínio de codificação usado no capítulo 4, as estatísticas mostram que poderá haver ganhos de desempenho. Mas, devido as restrições de tempo não foi possível implementar o algoritmo neste trabalho.

Na terceira etapa o dicionário foi populado com a inclusão de blocos deslocados. Foi avaliado o uso de blocos deslocados de múltiplos de $1/2$, $1/4$ e $1/8$ do seu tamanho. Verificamos, através da curva taxa-distorção, que a codificação de imagens suaves e mistas tiveram melhor desempenho com blocos deslocados de múltiplos de $1/4$ do seu tamanho. Foi também realizada uma análise estatística do uso dos subdicionários, que mostrou que, neste caso, valeria a pena usar a segmentação no espaço de códigos. Entretanto, devido a restrições de tempo isto não foi implementado neste trabalho.

Duas importantes conclusões podem ser tiradas deste trabalho. A primeira é que é de um modo geral vantajoso aumentar a cardinalidade dos dicionários com padrões típicos, como foi o caso dos blocos deslocados. Deve-se notar que, apesar do uso dos blocos deslocados não ter redundado em uma melhora, também não houve piora significativa apesar de um grande aumento de cardinalidade, o que reforça a tese de que o MMP tende a se beneficiar de um aumento de cardinalidade dos dicionários. A segunda conclusão é motivada pelos bons resultados obtidos com a segmentação dos dicionários no espaço de códigos, que sugere que ainda há bastante espaço para se melhorar o desempenho através de métodos mais eficazes de codificação por entropia. Disto concluímos que uma linha que vale a pena ser seguida no futuro é a investigação de formas alternativas de codificação por entropia no MMP, e a implementação da segmentação no espaço de códigos para os blocos deslocados, rotacionados e transpostos.

Referências Bibliográficas

- [1] CARVALHO, M. B., Compression of Multidimensional Signals based on Recurrent Multiscale Patterns. Tese de doutorado, Programa de Engenharia Elétrica/COPPE/UFRJ, Rio de Janeiro, RJ, Março 2001.
- [2] HUFFMAN, D. A., “A Method for the Construction of Minimum Redundancy Codes”. In: Proceedings of the IRE, v. 40, pp. 1098–1101, 1951.
- [3] SAYOOD, K., Introduction to Data Compression. 2 ed. San Francisco, Morgan Kaufmann Publishers, 2000.
- [4] ABRAMSON, N., Information Theory and Coding. New York, McGraw-Hill, 1963.
- [5] ZIV, J., LEMPEL, A., “A Universal Algorithm for Data Compression”. In: IEEE Trans. Inf. Theory, v. 23, pp. 337–343, 1977.
- [6] ZIV, J., LEMPEL, A., “A Compression of Individual Sequences Via Variable-Rate Coding”. In: IEEE Trans. Inf. Theory, v. 24, pp. 530–536, 1978.
- [7] HUANG, J. Y., SCHULTHEISS, P. M., “Block Quantization of Correlated Gaussian Random Variables”. In: IEEE Transactions on Communication Systems, pp. 289–296, 1963.
- [8] MACWILLIAMS, F. J., SLOANE, N., The Theory of Error Correcting Codes. Amsterdam: North-Holland, 1977.
- [9] CLARKE, R. J., Digital Compression of Still Images and Video. Academic Press, 1995.
- [10] TOPIWALA, P. N., Wavelet Image and Video Compression. 1 ed. Norwell, Kluwer Academic Publishers, 1998.

- [11] VETTERLI, M., KOVACEVIC, J., Wavelets and Subband Coding. Prentice Hall, 1995.
- [12] STRANG, G., NGUYEN, T., Wavelets and Filter Banks. 1 ed. Wellesley-Cambridge Press, 1996.
- [13] WITTEN, I., NEAL, R., CLEARY, J. G., “Arithmetic Coding for Data Compression”, Comm. ACM, v. 30, pp. 520–540, June 1987.
- [14] PENNENBAKER, W. B., MITCHELL, J. L., JPEG Still Image Data Compression Standard. 1 ed. Van Nostrand Reinhold, 1993.
- [15] SHANNON, C. E., “A Mathematical Theory of Communication”, Bell Syst. Tech. Journal, v. 27, 1948.
- [16] JAIN, A. K., Fundamentals of Digital Image Processing. Prentice Hall, 1989.
- [17] COVER, T. A., THOMAS, J. A., Elements of Information Theory. John Wiley and Sons, 1991.
- [18] PAPOULIS, A., Probability Rondon Variables, and Stochastic Processes. 2 ed. McGraw-Hill, 1984.
- [19] PEEBLES, P. Z., Probability Random Variables, and Random Signal Principles. 4 ed. New York, McGraw-Hill, 2001.
- [20] CARVALHO, M. B., SILVA, E. B., FINAMORE, W. A., “Multidimensional Signals Compression Using Recurrent Patterns”. In: Elsevier, pp. 1559–1580, 1995.
- [21] FILHO, E. B. L., Compressão de Imagens Utilizando Recorrência de Padrões Multiescala com Critério de Continuidade Inter-blocos. Tese de mestrado, PEE/COPPE/UFRJ, Rio de Janeiro, RJ, Março 2004.
- [22] DENN, M. M., Optimization by Variational Methods. 1 ed. New York, McGraw-Hill, 1969.
- [23] EVERETT, H., “Generalized Lagrange Multiple Method for Solve Problems of Optimum Allocation Resouces”. In: Operation Research, v. 11, pp. 399–417, 1963.

- [24] WU, S. W., GERSHO, A., “Rate-Constrained Optimal Block-Adaptative Coding for Digital Tape Recording of HDTV”. In: IEEE Trans. on Circuits and System for Video Tech., pp. 100–112, 1991.
- [25] SULLIVAN, G. J., BAKER, B. L., “Efficient Quadtree Coding of Images and Video”. In: IEEE Trans. on Image Processing, v. 3, pp. 327–331, 1994.
- [26] KIANG, S. Z., BAKER, R. L., SULLIVAN, G. J., et al., “Recursive Optimal Prune With Applications to Tree Structured Vector Quantizers”. In: IEEE Trans. on Image Processing, v. 1, pp. 162–169, 1992.
- [27] SAID, A., PEARLMAN, W., “A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees”. In: IEEE Transactions on Circuits and Systems for Video Technology, v. 6, Chicago, Illinois, June 1996.
- [28] “<http://www.cipr.rpi.edu/research/SPIHT>”.
- [29] RABBANI, M., JONES, P. W., Digital Image Compression Techniques. 1 ed. Bellingham, SPIE Press, 1991.

Apêndice A

Imagens

Neste apêndice encontramos as imagens *aerial*, *barbara*, *baboon*, *bridge*, *f16*, *gold*, *lena*, *pp1205* e *pp1209* usadas como dados de entrada para as simulações.

A.1 Imagem *aerial*



Figura A.1: Imagem *aerial*

A.2 Imagem *Baboon*

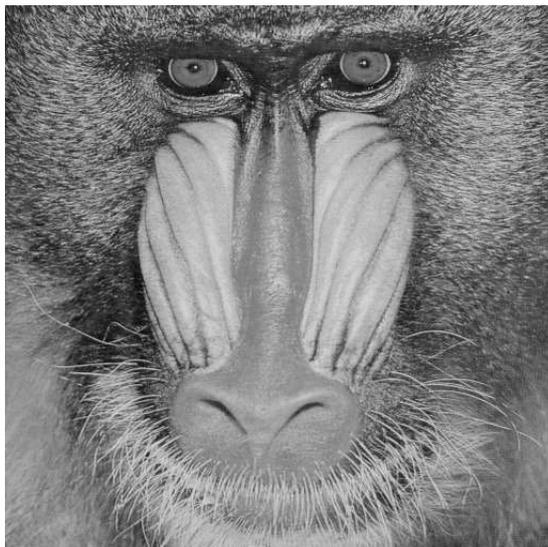


Figura A.2: Imagem *baboon*

A.3 Imagem *Barbara*



Figura A.3: Imagem *barbara*

A.4 Imagem *Bridge*



Figura A.4: Imagem *bridge*

A.5 Imagem *f16*



Figura A.5: Imagem *f16*

A.6 Imagem *gold*



Figura A.6: Imagem *gold*

A.7 Imagem *Lena*



Figura A.7: Imagem *lena*

A.8 Imagem pp1205

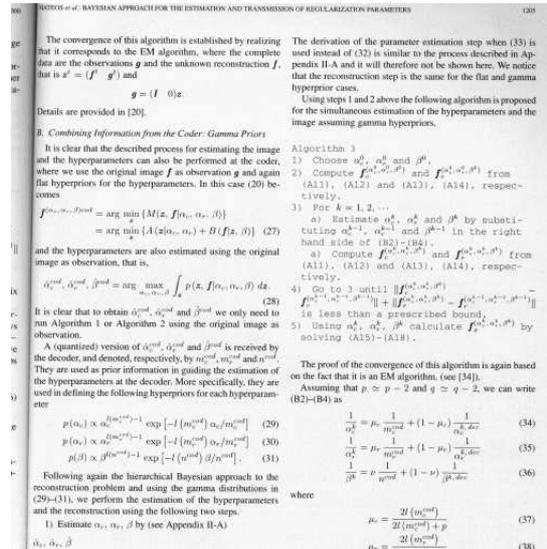


Figura A.8: Imagem pp1205

A.9 Imagem pp1209

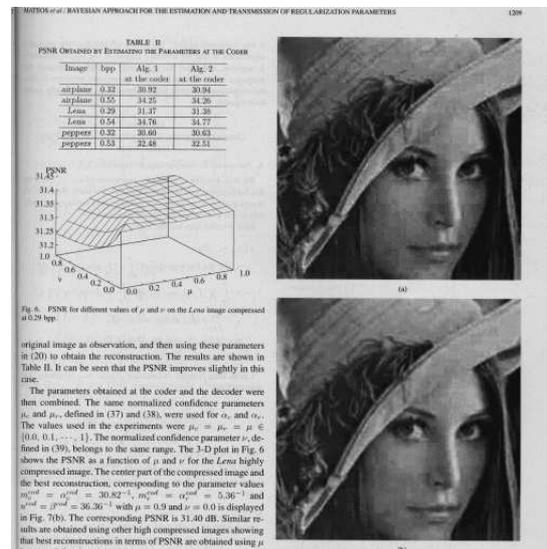


Figura A.9: Imagem pp1209