

3

Algoritmo Proposto para a Detecção de Cortes de Câmera

Este capítulo apresenta o algoritmo proposto para a detecção dos cortes de câmera, o qual procura atender aos diversos gêneros de vídeos e tipos de codificadores. São propostas heurísticas para identificação de padrões de comportamento das imagens e da codificação. A análise dos valores obtidos pelas métricas usuais revela a existência de padrões falsos de corte que estão simplesmente relacionados a escolhas feitas durante a história do processo de codificação do vídeo. Portanto, no algoritmo proposto, as análises das informações codificadas procuram identificar padrões estabelecidos pelo processo de codificação (a história da compactação), padrões esses considerados “assinaturas” dos codificadores. Para distinguir os quadros com características de corte real dos quadros com características influenciadas pelo codificador, são propostas filtragens para suavizar a influência destas assinaturas nos valores obtidos pelas métricas de caracterização de similaridade.

As métricas usuais para caracterização de semelhança a partir de informações do domínio comprimido, relacionadas no capítulo anterior, apresentam distribuições não ideais para os valores das análises dos quadros. A detecção de cortes de câmera é um problema com função de distribuição ideal desconhecida, ou seja, é desconhecido um critério de análise dos quadros para o qual exista pelo menos um limiar separando perfeitamente os quadros-corte dos quadros não-corte. Desta forma, não é possível garantir a existência de um único limiar para a classificação de todo o vídeo. As comparações quadro a quadro por limiar são extremamente dependentes da seqüência em análise e os cortes perdidos fogem a medidas estatísticas como a média e o desvio padrão. Por esse motivo, a abordagem proposta por esta pesquisa não está vinculada à busca de um limiar ótimo para uma função de caracterização dos cortes.

Neste capítulo, a Seção 3.1 apresenta as idéias para a busca por padrões de codificação e as seções restantes descrevem cada uma das etapas de refinamento e análise.

3.1. Busca de Padrão de Codificação

Considerando uma métrica qualquer para caracterização de corte com distribuição crescente, associando valores mais altos a quadros com as maiores probabilidades de corte, os possíveis valores assumidos por tal distribuição não-ideal podem ser classificados em até três faixas distintas (Figura 5):

- a primeira faixa inclui os valores mais altos que identificam apenas cortes verdadeiros;
- logo a seguir, a segunda faixa inclui valores de identificação de cortes verdadeiros e falsos;
- uma terceira faixa, contendo os valores mais baixos, inclui apenas quadros não-corte.

Entretanto, não há garantia da existência simultânea dessas faixas, qualquer que seja o vídeo ou a função de distribuição.

Métrica não ideal para caracterização de corte

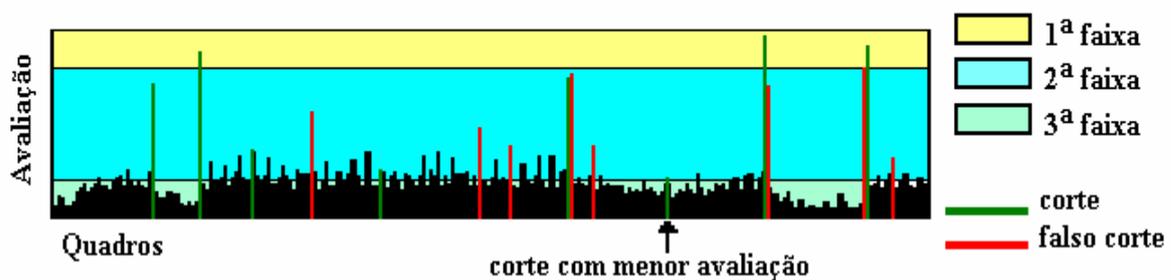


Figura 5 – Distribuição de uma métrica não ideal para caracterização de corte

Partindo desta classificação, proposta nesta dissertação, o presente trabalho propõe o uso de etapas sucessivas de refinamento do conjunto de quadros do vídeo para a identificação do subconjunto de quadros-corte.

Os refinamentos apresentados nas próximas seções usam critérios distintos para caracterização da similaridade entre os quadros. As métricas para análise da presença de corte são invertidas para detectar quadros semelhantes e os limiares escolhidos passam a pertencer a terceira faixa de valores das distribuições. Desta forma, as diversas etapas são reguladas pela escolha de limiares abaixo dos ideais, optando por eliminar os quadros abaixo do limiar, ao invés de classificar os

quadros com valores superiores. Assim, os quadros detectados como semelhantes são eliminados do conjunto de quadros candidatos aos cortes.

O padrão MPEG não estabelece a implementação de algoritmos de compressão para atender aos seus requisitos. A definição de procedimentos para atender aos seus requisitos fica a encargo da aplicação de compressão. Deste modo, o algoritmo de pareamento dos blocos entre quadros intercodificados e intracodificados, a distância máxima de procura dos Vetores de Movimento, a escolha do número de quadros nos GOPs, entre outras escolhas, são estabelecidas pelos mecanismos que definem o que, nessa dissertação, passa a ser denominado de “inteligência do codificador”. Sendo assim, diferentes implementações de compactadores MPEG, geram fluxos de dados (*streams*) distintos para a codificação de um mesmo vídeo.

Ao avaliar os primeiros resultados obtidos na busca pelas tomadas de câmera, observa-se que a análise das informações do domínio comprimido para detecção dos cortes de tomada de câmera necessita distinguir os comportamentos padrão de codificação (que, nesta dissertação, passam a ser chamados de “assinatura” do codificador) dos comportamentos indicativos da vizinhança de cortes de tomada de câmera. Desta maneira, é possível evitar erros de detecção associados a padrões de codificação.

A busca por essa distinção resulta na observação do comportamento dos codificadores em relação à propagação do erro de codificação existente entre os quadros P a P, no interior de um GOP. A análise desse comportamento está vinculada à sua interferência nas análises das taxas de referências entre os quadros.

Outras assinaturas identificadas por este trabalho foram encontradas pela análise da detecção de padrões distintos para o número de quadros no interior do GOP. Esses padrões são caracterizados pela distribuição dos tipos de quadros e pela distribuição dos Macroblocks nas imagens intercodificadas, decorrentes de sua posição em relação aos quadros intracodificados.

Alguns codificadores, por exemplo, forçam os quadros B, pertencentes ao final de GOPs, a referenciar o quadro I do próximo âncora, influenciando a decisão da referência em relação à proximidade aos quadros intracodificados. Este procedimento é feito na tentativa de diminuir o erro da imagem resultante, pela referência à imagem intracodificada mais próxima.

Uma vez identificada a existência de comportamentos específicos de grupos de codificadores, esta pesquisa estabeleceu, como requisito ao algoritmo desenvolvido, a adoção de critérios e métricas genéricas, ou seja, ser genérico quanto à natureza dos vídeos, sem estar vinculado a um compactador nem a modelos específicos de compactação. Em decorrência desse requisito, propõe-se um algoritmo desvinculado de um tamanho fixo de GOP, que não considera a existência de uma taxa ou padrão fixo de IPB-ratio e que fornece parâmetros de adaptação dos resultados a critérios de maior ou menor sensibilidade.

A escolha das métricas a serem usadas nos processos de refinamento é influenciada pela adoção de dois requisitos estabelecidos para garantir o desempenho do algoritmo:

- analisar apenas informações possíveis de serem extraídas diretamente do vídeo comprimido, sem a aplicação de processos de descompressão; e
- usar análises hierárquicas para descarte de grupos de quadros.

As informações consideradas para a análise dos cortes de câmera nas etapas de refinamento são as seguintes:

- (1) a distribuição dos tipos de quadros no interior dos GOPs;
- (2) o número de quadros por GOP;
- (3) o tamanho de armazenamento no fluxo de dados (*stream*) codificado dos GOPs e dos quadros intracodificados;
- (4) as taxas de RR e de Probabilidade de Corte Adaptada, por adaptação do algoritmo de Chen & Kuo (2002);
- (5) os coeficientes DC dos quadros I.

Além dessas informações, são fornecidos dois parâmetros ajustáveis:

- uma taxa de sensibilidade; e
- uma taxa de vizinhança mínima entre dois cortes.

O algoritmo proposto está baseado em etapas direcionadas ao descarte dos quadros considerados semelhantes, através da análise das informações no domínio comprimido, em oposição às abordagens de identificação direta de características dos cortes. Essas etapas procuram distinguir os padrões de corte de padrões considerados assinatura do codificador. A solução do conjunto de cortes é composta pela união dos conjuntos de quadros que não satisfazem a nenhuma das análises de semelhança.

Considerando o conjunto de quadros de um vídeo, as etapas de refinamento eliminam quadros em subconjuntos desse universo. Cada etapa de refinamento possui uma métrica para caracterização de semelhança entre quadros e uma regra de eliminação de quadros semelhantes. O conjunto solução é composto pelo agrupamento dos quadros-cortes. Um quadro é classificado como quadro-corte caso não seja eliminado por nenhuma das etapas de refinamento, logo, o conjunto solução é composto pela união dos quadros não eliminados pelas etapas de refinamento. A Figura 6 ilustra essas definições.

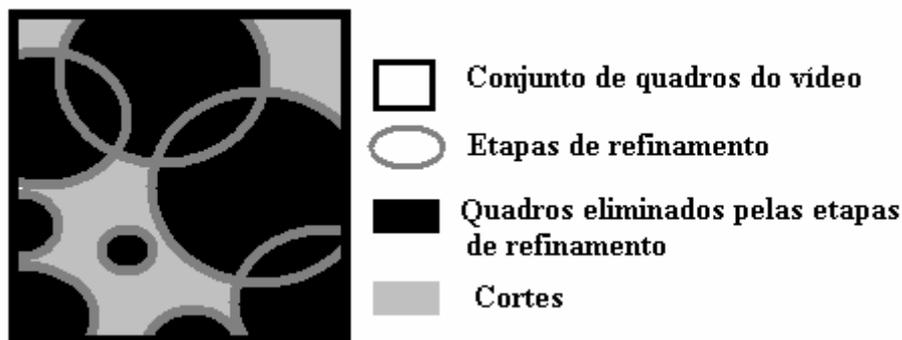
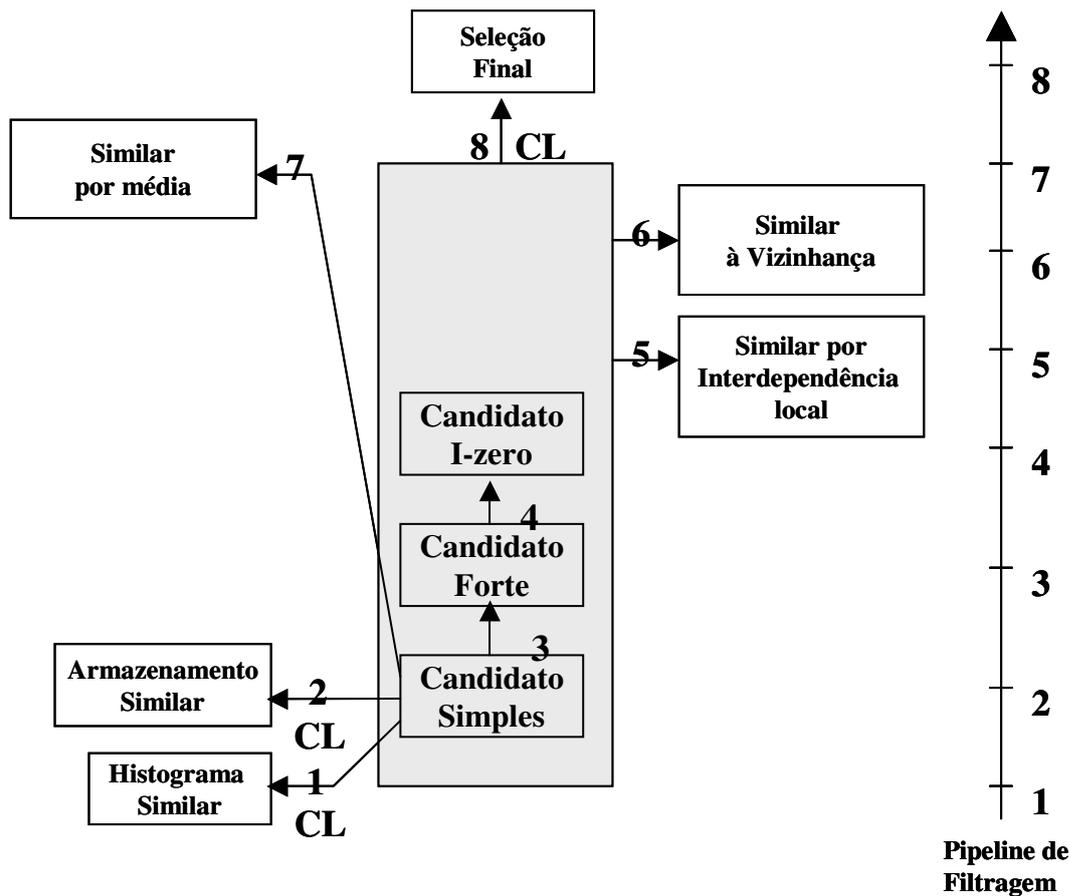


Figura 6 – Universo da segmentação

Durante todo o algoritmo, cada quadro possui um estado que representa a condição de sua candidatura como corte de câmera, conforme a Figura 7 abaixo. Todos os quadros do vídeo são inicialmente considerados candidatos aos cortes, recebendo a classificação de Candidatos Simples. Durante as etapas de refinamento, os candidatos podem ter sua candidatura promovida a Candidatos Fortes ou Candidatos I-zero. Os quadros eliminados em cada etapa têm sua candidatura desclassificada e não se tornam novamente candidatos, estabelecendo um fluxo em uma só direção.



***CL: classificação local**

Figura 7 – Estado da candidatura dos quadros

Nas próximas seções são descritas as etapas de refinamento e as respectivas transições de estados dos quadros.

3.2. Refinamento por Histograma

A primeira etapa de refinamento é feita por uma adaptação do algoritmo proposto por O'Toole (1998), para análise dos histogramas das imagens-DC do vídeo, apresentado no capítulo anterior. A escolha dessa métrica foi motivada por ser ela uma medida insensível ao movimento cromático no interior do quadro.

A obtenção das imagens-DC a partir de quadros P e B é feita por um processo de semidescompressão para recuperação dos coeficientes DC (Figura 8). Esses coeficientes são obtidos pela soma dos erros residuais, armazenados nos Macroblocks dos quadros P e B dependentes, com as interseções dos Macroblocks de referência. O quadro âncora referenciado, por sua vez, pode ser um quadro P,

fazendo que com essa etapa resulte em possíveis desdobramentos até o quadro I anterior.

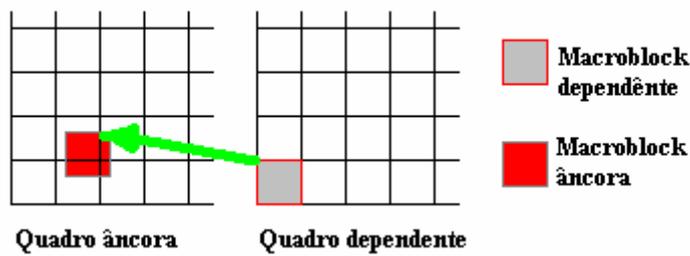


Figura 8 – Macroblocks dependentes

Nesta pesquisa, foram destacados os requisitos de eficiência, sendo para isso escolhidos critérios de análise com o menor custo possível, classificando as informações dos coeficientes DC das imagens P e B de obtenção onerosa. Por esse motivo, o algoritmo baseado na análise de histogramas, proposto para essa etapa de refinamento, analisa apenas a informação dos coeficientes DC dos quadros I. A análise entre quadros vizinhos é adaptada para a comparação entre os quadros I mais próximos. Nesse tipo de quadro, os coeficientes DC podem ser obtidos diretamente dos dados codificados, sem o processamento de reversão da predição nem da compensação de movimento.

Essa comparação é usada como critério hierárquico para análise do vídeo, permitindo a eliminação de grupos de quadros. Duas imagens I com histogramas considerados semelhantes indicam a inexistência de um corte de câmera entre elas (a menos que ocorra uma tomada de câmera completa no interior do intervalo entre as duas imagens I, seguida pela tomada de câmera inicial). Além de ser extremamente raro, esse evento pode ser captado pelo aumento da sensibilidade do algoritmo.

A análise hierárquica é vantajosa também por permitir a eliminação dos falsos cortes causados por flashes, superexposições a fontes luminosas, explosões ou efeitos pisca-pisca, caracterizados pela alteração do conteúdo de poucos quadros, seguidos de imagens semelhantes às anteriores a esses eventos.

A comparação dos histogramas das imagens I é feita pela análise do ângulo entre eles, encontrado pela CSM. Uma particularidade dessa medida é que considerando cada nível de cinza como uma faixa de intensidade, faixas vizinhas influenciam na análise com a mesma intensidade de faixas distantes. Essa

igualdade de pesos não considera as características de percepção visual intrínsecas à detecção de cortes de câmera.

Como forma de suavizar a contribuição dos níveis de intensidade muito próximos, foram feitos agrupamentos desses níveis em uma mesma faixa, procurando transmitir sua não distinção perceptível. Quanto maior a quantidade de faixas, maior a sensibilidade a pequenas mudanças de intensidade entre as imagens, ou seja, mais imagens são consideradas distintas. Ao diminuir o número de faixas e, portanto, as dimensões do histograma, mais imagens são consideradas semelhantes, tendendo a surgir mais imagens com ângulo zero entre seus histogramas.

O extremo acontece ao agrupar todos os tons de intensidade do histograma em apenas uma faixa, descaracterizando totalmente a análise da semelhança entre as duas imagens. Faz-se necessário um equilíbrio entre o número de faixas utilizadas para a caracterização do problema e o tamanho da vizinhança dos níveis de intensidade a serem considerados semelhantes em relação à percepção visual.

Considerando não ser do interesse da detecção de tomadas de câmera a distinção entre tons de intensidade muito próximos, é proposto o agrupamento de alguns tons em uma mesma faixa do histograma. Como resultado obtém-se uma menor distinção entre os vetores de imagens com intensidades próximas.

Calculados os ângulos de comparação entre os quadros I consecutivos de todo o vídeo, esses valores são ordenados usando uma estrutura de árvore de ordenação. Essa estrutura foi adotada por esta pesquisa para permitir a comparação dos valores assumidos na classificação de uma determinada característica em relação a todo o vídeo, por uma ordenação crescente. Desta forma, é possível separar para o descarte, os quadros que recebem as menores avaliações. Essa organização também permite identificar a irrelevância de um determinado critério, em relação a um vídeo específico, quando os valores assumidos pelos quadros apresentam-se agrupados em poucos nós da árvore, sem diferenciações significativas. Nesse caso, o critério de avaliação é desconsiderado, aumentando a maleabilidade do algoritmo a vídeos genéricos, ao deixar de eliminar cortes erroneamente, pela análise de um comportamento inexistente.

Obtida a separação dos quadros I com menor valor de histograma do vídeo, esses quadros passam a ser considerados semelhantes ao quadro I anterior correspondente, de forma a terem sua candidatura a corte eliminada.

A diferença do espaço de armazenamento entre GOPs vizinhos é calculada de forma ponderada, permitindo a comparação posterior, na mesma árvore de ordenação, de trechos do vídeo com imagens de natureza distinta. A Diferença Ponderada de Armazenamento de GOPs (DifPG) entre os GOPs G1 e G2 é definida pela equação:

$$DifPG(G1, G2) = \frac{|tamanho(G1) - tamanho(G2)|}{|tamanho(G1) + tamanho(G2)|} \quad (5)$$

Onde:

- **G1 e G2** representam GOPs consecutivos analisados.

A comparação dos tamanhos de armazenamento de GOPs vizinhos só é feita entre GOPs com o mesmo número de quadros. As diferenças de armazenamento entre GOPs vizinhos de mesmo tamanho em todo o vídeo são inseridas em uma árvore de ordenação. A partir dessa ordenação, os GOPs com menor diferença em relação a seus vizinhos anteriores são considerados similares. Neste caso, os quadros no intervalo entre o quadro I do segundo GOP e o último quadro âncora são descartados e passam para o estado “Armazenamento Similar” (Figura 11). Esse intervalo estabelece a fronteira máxima de referências no interior do segundo GOP, já que os quadros posteriores ao último quadro âncora podem referenciar quadros do GOP seguinte.

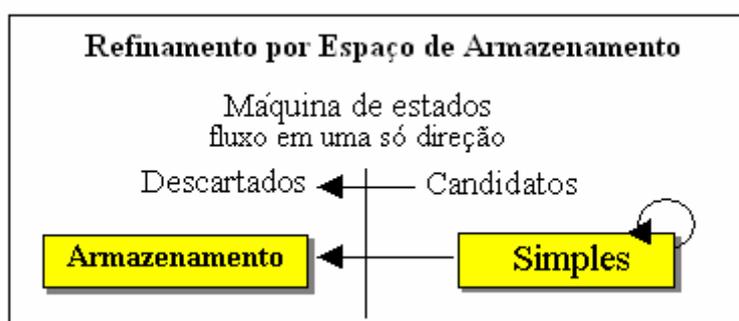


Figura 11 – Mudança de estado da candidatura dos quadros (Refinamento por Espaço de armazenamento).

A comparação do espaço de armazenamento dos GOPs busca uma assinatura do processo de compressão. Os GOPs que ocupam espaços aproximadamente iguais são comprimidos com um mesmo padrão de compressão. O espaço ocupado é preenchido principalmente pela codificação dos quadros

âncoras, tendo portanto, um maior peso no espaço total encontrado. Dessa forma, ao encontrar diferenças desprezíveis entre GOPs vizinhos, a distribuição das referências entre os quadros âncoras se mantém, não existindo um evento estranho entre o primeiro quadro I de cada GOP e seu último quadro P. Essa distribuição está relacionada à correção da propagação do erro de codificação através das imagens P do GOP.

Paralelamente à análise do espaço ocupado pelos GOPs, é feita uma análise do espaço ocupado pelos quadros I. Cabe ressaltar que as taxas de compressão das imagens intracodificadas, traduzidas no tamanho de seu armazenamento, são relacionadas com a complexidade das frequências de seu conteúdo. Espera-se que dois quadros I com o mesmo conteúdo sejam codificados ocupando espaços similares.

A diferença ponderada entre os espaços ocupados pelos quadros I mais próximos é fornecida pela Equação 6 (Diferença Ponderada de Armazenamento entre Quadros - DifPQ):

$$DifPQ(I1, I2) = \frac{|tamanho(I1) - tamanho(I2)|}{tamanho(I1) + tamanho(I2)} \quad (6)$$

Onde:

- **I1** e **I2** representam quadros intracodificados mais próximos.

As diferenças ponderadas dos espaços ocupados pelos quadros I de todo o vídeo são ordenadas. Os quadros I com menores diferenças são considerados similares por essa característica. Nesse caso, é feita a desclassificação dos quadros no intervalo entre o primeiro e o segundo quadro I, este incluído. Os quadros desclassificados possuem todas as suas referências relativas aos dois quadros I considerados similares.

3.4

Refinamento pelo Padrão de Distribuição dos Quadros

A etapa de refinamento a seguir procura alterações na assinatura do codificador em relação à ordem e ao tipo das imagens no interior dos GOPs. Essa assinatura se caracteriza pela confirmação da existência de uma distribuição

ordenada dos quadros I, P e B, repetida na maioria dos GOPs do vídeo e de um número de quadros por GOP padrão.

Identificada uma assinatura relacionando o posicionamento mais comum dos tipos de quadros no interior do GOP, os quadros que fogem a essa assinatura são considerados indicativos da vizinhança de corte, como causa da mudança de codificação, fugindo ao padrão da assinatura.

Essa etapa não desclassifica quadros, mas estabelece uma proteção maior à candidatura de quadros que não foram eliminados pelas etapas anteriores e que fogem ao padrão da assinatura encontrada (Figura 12).

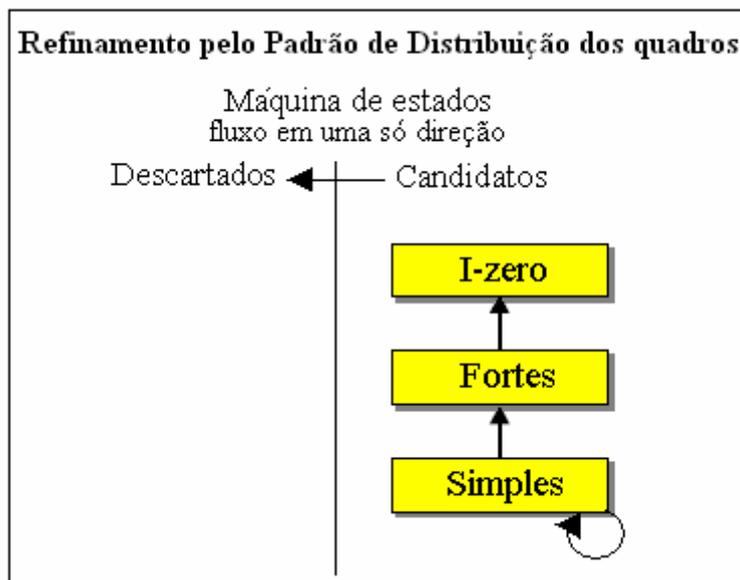


Figura 12 – Mudança de estado da candidatura dos quadros (Refinamento por Padrão de Distribuição dos quadros).

A primeira forma de promoção da candidatura dos quadros eleva aqueles codificados com tipos que fogem ao padrão esperado em relação à sua posição no GOP. Esses quadros passam a ser considerados Candidatos Fortes.

A intenção de promover a candidatura desses quadros está baseada na suposição de que a alteração do tipo do quadro está relacionada ao aumento do erro de codificação na vizinhança dos cortes.

A alteração do tipo do primeiro quadro no interior do GOP, quadro na posição zero, para um quadro do tipo I, é tratada de forma especial. Essa alteração de padrão é comum a vídeos que foram comprimidos usando segmentação em tomadas de câmera. Inicialmente, o codificador identifica as fronteiras de tomadas

de câmera e codifica essas imagens como intracodificadas. Esse procedimento ocorre com o intuito de otimizar a compressão.

O mesmo acontece nos limites de edição de vídeos formados pela simples junção de outros vídeos pré-compactados sem a aplicação de procedimentos de ajuste e recompactação dessas bordas. Os vídeos originais utilizados na edição iniciam-se em quadros I, deixando um rastro onde essas fronteiras aparecem. Por pertencer a duas tomadas de câmera distintas, essas fronteiras representam posições de corte de câmera.

Seguindo a análise da assinatura, ao detectar essa alteração de padrão em candidatos posicionados no primeiro quadro do GOP, o candidato é promovido ao estado mais significativo da classificação dos quadros, denominada Candidato I-zero.

Uma ressalva deve ser feita em relação a vídeos com assinatura que identificam quadros intracodificados na posição zero do GOP. Deixando de ser uma exceção do comportamento do codificador, esses quadros perdem sua relevância e o estado de suas candidaturas é então mantido. Esse tipo de vídeo não possui Candidatos I-zero.

3.4. Probabilidade de corte

As próximas etapas de refinamento estão baseadas na análise do comportamento das interdependências entre quadros vizinhos pela análise das referências de seus Macroblocks.

O uso de máscaras para o cálculo da Função de Probabilidade proposta por Chen & Kuo (2002), limita a aplicação de seu algoritmo a vídeos de IBP-ratio constantes. A partir dessa proposta do uso de máscaras de tamanho fixo durante todo o vídeo, sugere-se, na presente dissertação, uma generalização para a divisão em subsequências. Uma subsequência é definida por um conjunto de zero ou mais quadros B, obrigatoriamente seguido de um quadro âncora (Figura 13). Essa definição estabelece uma partição dos GOP, agrupando os quadros por áreas de influência de suas referências, em relação à possibilidade de corte .

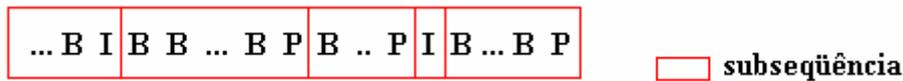


Figura 13 – Separação em subseqüências

O cálculo da probabilidade de corte é, então, feito por subseqüências, adaptando a função de probabilidade, para considerar distribuições diversas durante um mesmo vídeo. Esta adaptação proposta por esta dissertação é denominada Função de Probabilidade Adaptada e é definida por:

$$Pr ob. Adaptada(i) = \left(1 - \frac{\sum_{i=0}^{c-1} BR_i^2 + \sum_{i=c}^{i=n} FR_i^2}{\sum_{i=0}^{c-1} BR_i^2 + \sum_{i=c}^{i=n} FR_i^2}\right) \quad onde 0 \leq i \leq n \quad (7)$$

Outra modificação proposta ao trabalho de Chen & Kuo (2002) está relacionada aos Macroblocks Skipped e Bidirecionais. Não foram encontradas na literatura observações sobre o tratamento dos Macroblocks Skipped. Nesta pesquisa, sua análise é feita somando a percentagem desses Macroblocks ao valor de taxa de referência (RR) da direção equivalente à sua referência, por serem uma cópia exata do Macroblock referenciado.

Os Macroblocks Bidirecionais, por sua vez, são usualmente tratados da mesma forma que os de referência para trás (*Backward*). Entretanto, essa consideração não estabelece nenhuma relação de semelhança desses Macroblocks em relação à direção para frente (*Forward*). A possibilidade de descartar a análise desse tipo de Macroblock também não estabelece a similaridade existente entre os dois quadros referenciados por esse tipo de Macroblock. Desta forma, é proposta a participação dos Macroblocks Bidirecionais, no cálculo da Função de Probabilidade Adaptada, com um fator de ponderação de 50%, para representar a relação de ponderação entre as imagens referenciadas e estabelecer uma participação no cálculo da probabilidade com menor intensidade do que as referências unidirecionais.

A análise da distribuição da Função de Probabilidade Adaptada resulta na descoberta de padrões de distribuição relacionados com características do

processo de compreensão. Essa assinatura do codificador está vinculada aos tipos de quadros e sua posição relativa ao interior do GOP, neste texto denominados “canal de posição”.

As Figuras 14, 15 e 16 apresentam as distribuições de três vídeos, com gênero, parâmetros de codificação, e codificador distintos, para ilustrar o comportamento por canal da distribuição da Função de Probabilidade Adaptada. A observação do comportamento destas distribuições e o surgimento de falsos picos de probabilidade revelam a periodicidade da Função de Probabilidade Adaptada, mascarando a caracterização dos quadros-cortes.

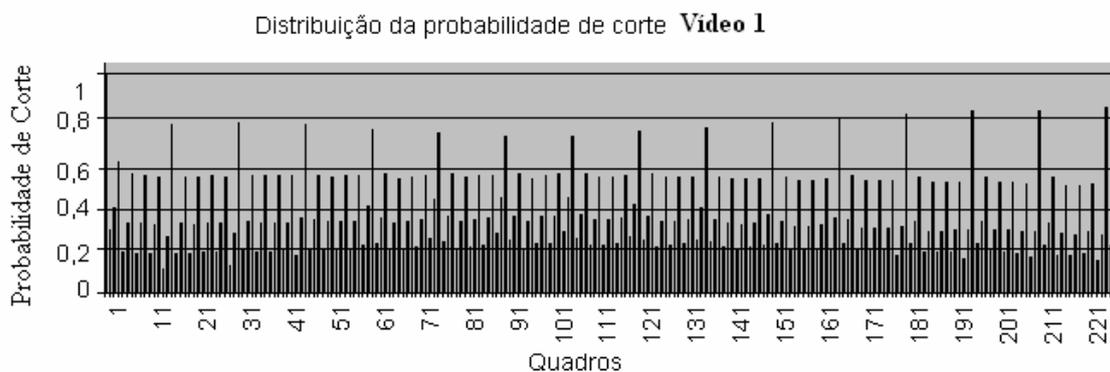


Figura 14 – Distribuição da Função de Probabilidade Adaptada – Vídeo 1

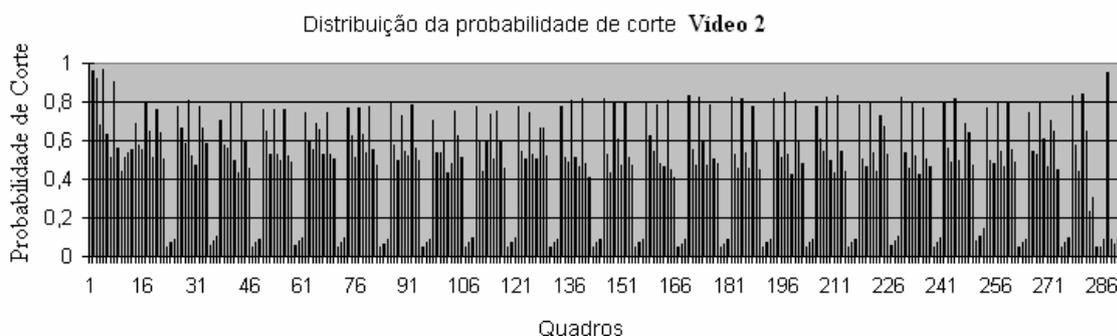


Figura 15 – Distribuição da Função de Probabilidade Adaptada – Vídeo 2

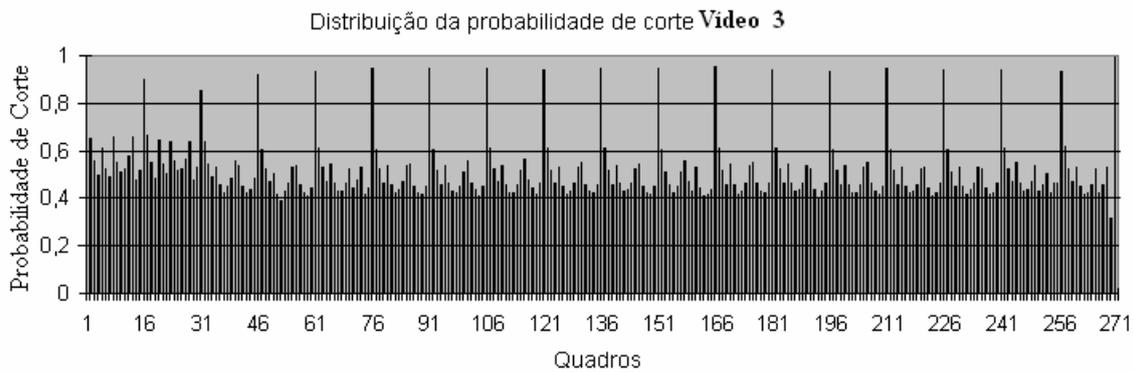


Figura 16 – Distribuição da Função de Probabilidade Adaptada – Vídeo 3

Essas características foram comprovadas pela análise da média e do desvio padrão da distribuição da Função de Probabilidade Adaptada, tomadas em relação a cada canal, demonstrando a existência de uma influência aos valores assumidos pela distribuição. Essa influência é inerente ao processo de compactação pela correção do erro de codificação, propagado através dos quadros interdependentes. Por esse motivo, a média e o desvio padrão, em relação a cada canal, foram adotados como uma métrica do comportamento padrão do codificador, sendo calculados a cada vídeo.

Como forma de identificar o significado real da avaliação fornecida pela Função de Probabilidade Adaptada, as análises feitas pelas etapas de refinamento seguintes usam uma ponderação dos valores obtidos por essa métrica em relação ao canal ocupado pelo quadro em análise. Dessa forma, busca-se distinguir a natureza real das referências relacionadas ao aparecimento de cortes, de referências relativas a comportamentos padrões do processo de codificação.

3.5.

Definição da Probabilidade Ponderada por Canal:

A Probabilidade Ponderada por Canal (ou Probabilidade Ponderada) é dada pela seguinte Equação:

$$Pr obabilidadePonderada (i) = \frac{Pr obabilidade Adaptada (i) - média (c)}{desvio padrão (c)} \quad (8)$$

Onde:

- **i** representa o quadro analisado;

- c representa o canal ocupado pelo quadro no interior do GOP.

A Figura 17 auxilia a compreensão dos índices envolvidos no cálculo da Probabilidade Ponderada.

A comparação para obter a probabilidade máxima ponderada entre duas probabilidades nos casos em que, pelo menos, um dos quadros esteja em um canal de desvio padrão igual a zero, é feita pela comparação dos valores obtido na Função de Probabilidade Adaptada.

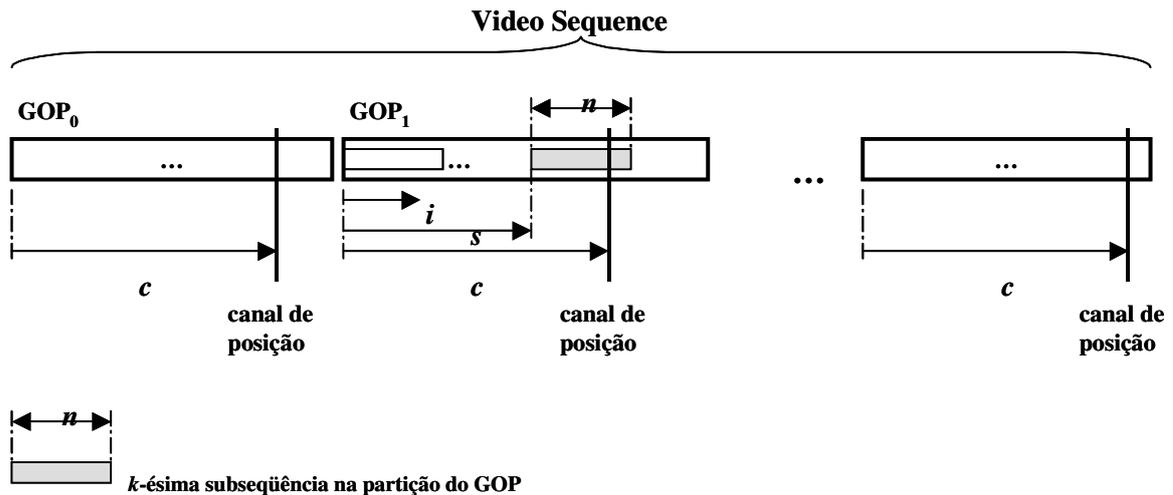


Figura 17 – Canal de posição

Tendo definido os procedimentos de ponderação para distinguir essa assinatura de codificação, as etapas seguintes propõem o descarte dos quadros nos vales e depressões da Função de Probabilidade Adaptada.

3.6. Refinamento pela Vizinhança no interior das subseqüências

A primeira análise da Função de Probabilidade Adaptada é uma comparação entre quadros vizinhos no interior de uma mesma subseqüência (Figura 18).

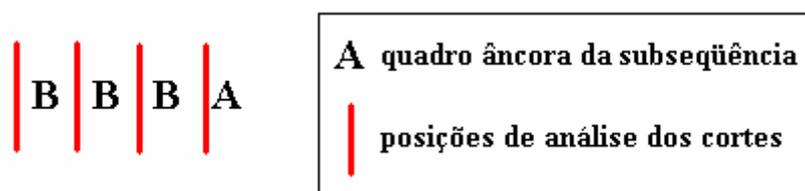


Figura 18 – Possibilidade de corte no interior de uma subseqüência

As taxas de referência envolvidas nos cálculos das probabilidades de dois quadros vizinhos no interior da mesma subsequência são todas iguais, com exceção da taxa de referência do primeiro desses dois quadros comparados. A possibilidade de corte anterior ao primeiro quadro é avaliada usando sua taxa de referência para frente (*Forward*), enquanto a possibilidade de existência de corte após esse quadro é analisada usando sua taxa de referência para trás (*Backward*).

A comparação entre quadros vizinhos usando apenas essas taxas alteradas permite que a busca pela existência de corte entre esses quadros seja simplificada para uma comparação entre o valor da taxa de referência para frente (*Forward*) ou para trás (*Backward*) do primeiro dos dois quadros comparados. A taxa de referência para frente (*Forward*) é um indicativo de semelhança aos quadros anteriores, enquanto que a taxa de referência para trás (*Backward*) é um indicativo de semelhança aos quadros posteriores.

Caso a taxa de referência para frente (*Forward*) do primeiro quadro comparado seja maior que sua taxa de referência para trás (*Backward*), a possibilidade de corte é maior para o para o segundo quadro da comparação, logo, o primeiro quadro pode ser eliminado. No caso oposto, em que a taxa de referência para trás (*Backward*) é maior que sua taxa de referência para frente (*Forward*), a possibilidade de corte é maior para o primeiro quadro da comparação, logo, o segundo quadro comparado é eliminado.

A partir das observações acima, é feita a eliminação de um a cada dois quadros consecutivos no interior de uma subsequência pela comparação de semelhança ao seu vizinho. Por extensão, em relação ao interior da subsequência, o quadro mantido deve possuir probabilidade maior do que a do quadro seguinte e também maior do que a probabilidade do quadro anterior.

Esse descarte procura picos em relação aos dois vizinhos mais próximos, mas deve, também, levar em consideração a influência da assinatura do codificador para que quadros em canais de probabilidade média alta não descartem seus vizinhos indiscriminadamente.

Dessa forma, os candidatos com probabilidade ponderada menor que a de seus vizinhos mais próximos no interior da subsequência são desclassificados, passando para o estado de “Crescimento” (Figura 19).

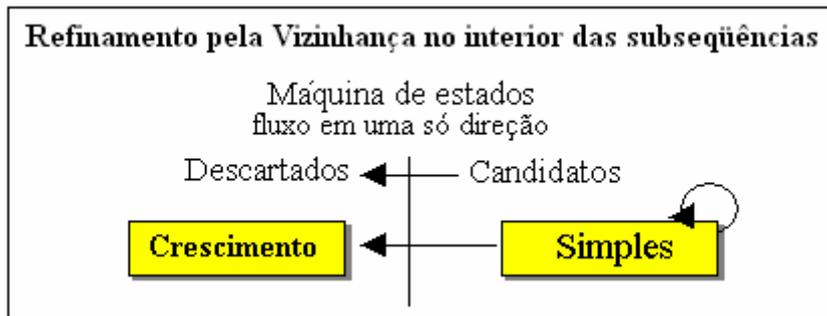


Figura 19 – Mudança de estado da candidatura dos quadros (Refinamento pela Vizinhança no interior das subseqüências).

3.7. Refinamento por Vizinhança Mínima

O tempo percorrido entre dois cortes de câmera apresenta variações significativas de acordo com o gênero do filme examinado. A linguagem visual utilizada em intervalos comerciais e filmes de ação, por exemplo, utiliza a apresentação de *flashes* de tomadas de câmera em um pequeno intervalo de tempo. Vídeos de documentários, noticiários, filmes de drama e romance tendem a apresentar tomadas de câmera mais longas, com duração de alguns segundos.

Essa característica semântica permite incorporar mais um processo de refinamento ao conjunto de candidatos. A análise da distância entre os candidatos permite a eliminação de candidatos muito próximos, interiores a um intervalo considerado mínimo entre dois cortes.

O refinamento por Vizinhança Mínima, proposto nesta dissertação, busca a adaptação do conjunto de candidatos para um conjunto de máximos locais em relação aos valores de probabilidade ponderada, procurando descartar os quadros que não são máximos ponderados usando a vizinhança mínima como janela de análise. Esse refinamento está vinculado a uma escolha inicial de segmentação do vídeo com espaçamento entre cortes, em intervalos maiores que uma determinada fração de segundo pré-definida. Essa distinção intuitiva permite o ajuste para a aceitação de movimentos rápidos sem a detecção de falsos cortes de câmera.

Nessa etapa, ao encontrar dois candidatos no interior de uma vizinhança mínima esperada, um dos dois quadros é eliminado. O primeiro critério da comparação dos candidatos no interior da vizinhança mínima é a existência de um

Candidato I-zero. Nesse caso, o quadro de candidatura Simples ou Forte é desclassificado por vizinhança e o candidato I-zero mantém o seu estado.

O segundo critério de comparação é usado para avaliar candidatos Simples e Fortes. Segundo esse critério, o quadro que possui probabilidade máxima ponderada mantém seu estado e o outro é desclassificado por “Vizinhança” (Figura 20).

Um terceiro critério é usado para o desempate. No caso de empate, a candidatura do primeiro quadro da comparação é mantida e, conseqüentemente, o segundo é eliminado. Considerando que ambos pertencem à mesma tomada de câmara, esse critério assegura maior proximidade à fronteira do corte.

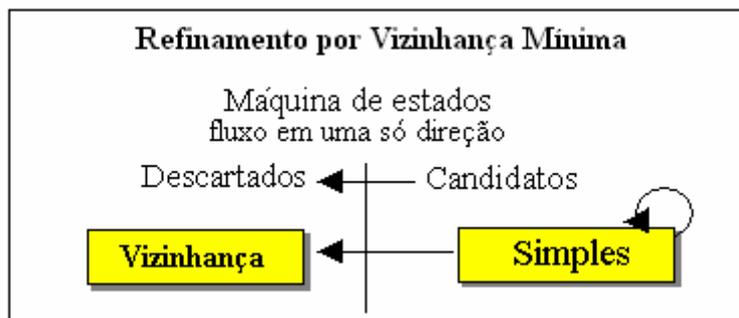


Figura 20 – Mudança de estado da candidatura dos quadros (Refinamento por Vizinhança Mínima).

As taxas de exibição mais comuns variam em torno de 30 quadros por segundo. Considerando esta taxa de quadros por segundo e três opções de segmentação, um segundo, meio segundo, ou sem eliminação por vizinhança, esses valores são traduzidos para este refinamento como uma vizinhança mínima no interior de 30 quadros, 15 quadros, ou sem esse refinamento, respectivamente.

3.8. Refinamento dos Candidatos Simples por Canal

Nessa etapa, é feita uma ordenação dos Candidatos Simples usando uma nova Árvore de Ordenação. A análise direta desses quadros, em relação aos valores da Função de Probabilidade Adaptada de corte, não leva em consideração a existência de uma assinatura na distribuição da probabilidade, por posição no interior do GOP. Usando esse critério, quadros posicionados em canais de probabilidade média alta são comumente ordenados à frente de quadros-corte,

posicionados em canais de probabilidade média baixa, levando a erros de detecção, pela não distinção da assinatura do codificador.

Para considerar a existência da assinatura do codificador, ou seja, existência de um padrão de referência dependente do posicionamento do quadro no interior do GOP, os valores repassados a árvore de ordenação representam a probabilidade ponderada dos quadros.

Apenas os quadros Candidatos Simples com probabilidade maior do que a probabilidade média do respectivo canal ocupado são inseridos na árvore de ordenação. Os demais Candidatos Simples são descartados pela justificativa de que as probabilidades assumidas caracterizam apenas o padrão do codificador para o canal ocupado, sem relevância para a detecção dos cortes de câmera. Esses quadros mudam para o estado denominado “Semelhança por Média” (Figura 21).

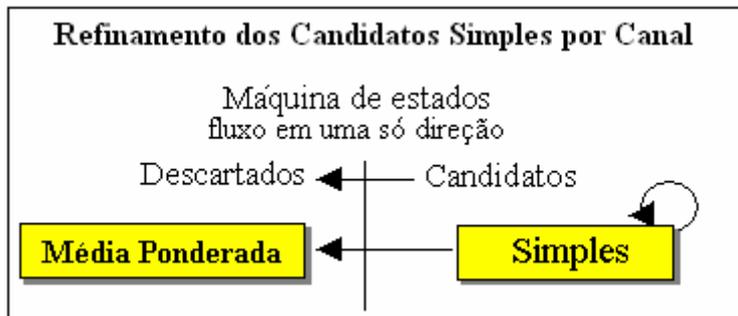


Figura 21– Mudança de estado da candidatura dos quadros (Refinamento dos Candidatos Simples por Canal).

Assim, os Candidatos Simples que não foram desclassificados pela média são inseridos na árvore de ordenação dos Candidatos Simples, usando como peso a sua probabilidade ponderada.

3.9. Seleção final

Nessa etapa é feita a seleção final do conjunto de quadros a serem classificados como corte. Os quadros que ainda não foram descartados são analisados de acordo com a classificação de sua candidatura.

Todos os quadros que finalizaram a análise no estado I-zero, ou seja, não foram eliminados pelas etapas anteriores e possuem, pela sua definição, probabilidade 1, são considerados cortes. O mesmo acontece aos quadros considerados Candidatos Fortes com probabilidade 1.

O parâmetro de sensibilidade, fornecido ao algoritmo, regula a participação dos Candidatos Fortes e Simples restantes. A partir desse parâmetro, é estabelecido um teto ao número de quadros considerados corte nessas categorias. Considerando esse limite, esses candidatos completam a lista de cortes, em ordem de probabilidade ponderada.