

## METAHEURÍSTICAS APLICADAS À MELHORIA DO DESEMPENHO DO H.264 DA TV DIGITAL BRASILEIRA

**Iris Correia das Chagas Link**

Universidade do vale do Rio dos Sinos - PIPCA  
Av. Unisinos, 950. São Leopoldo, RS. 93022-000  
[linck.iris@gmail.com](mailto:linck.iris@gmail.com)

**Arthur Tórgo Gómez**

Universidade do vale do Rio dos Sinos - PIPCA  
Av. Unisinos, 950. São Leopoldo, RS. 93022-000  
[atgomezbr@gmail.com](mailto:atgomezbr@gmail.com)

**Marta Becker Villamil**

Universidade do vale do Rio dos Sinos - PIPCA  
Av. Unisinos, 950. São Leopoldo, RS. 93022-000  
[mbVillamil@gmail.com](mailto:mbVillamil@gmail.com)

### RESUMO

Este trabalho apresenta uma abordagem utilizando metaheurísticas para o problema de configuração de parâmetros do codificador de vídeo H.264. O desempenho do H.264 depende da configuração de seus parâmetros. Devido à alta flexibilidade da configuração desses parâmetros, o H.264 pode ter um baixo desempenho se suas configurações não forem feitas de forma apropriada. Com o propósito de atenuar este problema, foi desenvolvido um algoritmo híbrido, o qual usa a Busca Tabu e o Algoritmo Genético, que simula o comportamento do codec de vídeo H.264. À medida que o H.264 tem seus parâmetros alterados pelo algoritmo, seu comportamento é avaliado por uma função objetivo. A melhor função objetivo será aquela que alcançar uma melhor configuração para os parâmetros do H.264. Somente 6 parâmetros foram considerados e estudados neste trabalho, os quais são: *bit rate*, *frame rate*, parâmetros de quantização dos *slices* tipo I, B e P e por fim, a quantidade de *slice* tipo B em um GOP (Group Of Picture – Grupo de figuras)

**PALAVARAS CHAVE.** Algoritmo Genético, Busca tabu, CODEC H.264/AVC.

**TEL&SI - PO em Telecomunicações e Sistemas de Informações**  
**MH - Metaheurísticas**

### ABSTRACT

This work presents an approach utilizing metaheuristics to the problem of H.264 parameters configuration. The H.264 performance depends of its configuration parameters. Due to the high flexibility of its configuration parameters, the H.264 can have a low performance if its setting is not done properly. In order to try to mitigate this problem, it was developed a hybrid algorithm, composed of the Tabu Search and the Genetic Algorithm, which simulates the H.264 video codec behavior used in the Brazilian Digital TV. As the H.264 has its parameter values changed by the algorithm, its behavior is evaluated by the objective function. The best objective function will be that achieve a better configuration for the H.264 parameters. Only six parameters were considered and studied in this work, that are: bit rate, frame rate, quantization parameters of I, B and P slices and lastly the number of B slice in a GOP (Group Of Picture).

**KEYWORDS.** Genetic Algorithm. Tabu Search. H.264 CODEC.

## 1. Introdução

No ano de 2005, o governo brasileiro deu início ao projeto do Sistema Brasileiro de TV Digital - SBTVD (ABNT, 2007) onde houve a participação de diversas universidades e centros de pesquisa. O projeto SBTVD aborda todas as partes que compõem um sistema de TV Digital e inovações que objetivaram ajustar a implantação da TV Digital à realidade dos brasileiros. A TV Digital Brasileira adotou como padrão de codificação de vídeo, o padrão H.264/MPEG-4 AVC (ITU-T, 2007), também conhecido como H.264. Este padrão possui inúmeros parâmetros de configuração que o deixam com alta flexibilidade e afetam enormemente o seu desempenho. Devido a isso, um codificador H.264 configurado inadequadamente pode apresentar um desempenho muito inferior à sua real capacidade.

No Sistema Brasileiro de Televisão Digital, o canal de transmissão adotado é o mesmo dos sistemas analógicos e tem uma largura de faixa de 6 MHz centrado em frequência das faixas VHF e UHF (CPqD, 2006a, 2006b). A codificação de sinais-fonte trata a questão da compressão dos sinais de áudio e vídeo transportados pelo sistema e dos métodos de transporte de dados. A codificação de sinais fonte é um dos principais viabilizadores da TV Digital, dada as taxas de bits relativamente elevadas, demandadas para transmissão destes sinais, e a necessidade de se estruturar os dados que a tecnologia permite transmitir (CPqD, 2006a, 2006b). Vários trabalhos foram feitos a fim de otimizar a performance do CODEC de vídeo H.264 (Cermak, Pinson, & Wolf, 2011; Correia & Pereira, 2003; Fiems, Steyaert, & Bruneel, 2012; Nemethova, Ries, & Rupp, 2004; Tianbing & Weidong, 2006; Yasakethu, Fernando, Adedoyin, & Kondo, 2008; Yu-Wen, Bing-Yu, Shao-Yi, Shyh-Yih, & Liang-Gee, 2006). Para avaliar o desempenho e a qualidade da imagem produzida por esses algoritmos são usadas métricas objetivas e subjetivas (ITU-R, 2002; Jianning, Yuwen, Shiqiang, & Yuzhuo, 2003; Malvar, Hallapuro, Karczewicz, & Kerofsky, 2003; Moriyoshi, Shinohara, Miyazaki, & Kuroda, 2001; Sikora, 2005).

Neste trabalho é apresentado um algoritmo híbrido que simula o comportamento do Codificador/Decodificador de vídeo H.264, ou simplesmente CODEC H.264, utilizado no Sistema Brasileiro de Televisão Digital. O algoritmo proposto, denominado Simulador de Metaheurísticas aplicado a um CODEC H.264, ou simplesmente SMAC, tem a finalidade de buscar a melhor configuração possível de seis dos principais parâmetros utilizados para a configuração do CODEC H.264. Este problema é abordado como um problema de otimização combinatória conhecido como Problema de Seleção de Partes e que é classificado como NP-Difícil. O algoritmo híbrido proposto, SMAC, foi desenvolvido com base em duas metaheurísticas: Busca Tabu e Algoritmo Genético. Os seis parâmetros de configuração a serem otimizados pelo algoritmo são: o *bit rate*; o *frame rate*; os parâmetros de quantização de quadros tipo B, tipo P e tipo I e a quantidade de quadros tipo B em um grupo de imagens (GOP – Group of Pictures). Os dois primeiros parâmetros mencionados atuam basicamente sobre a qualidade da imagem do vídeo enquanto que os demais parâmetros atuam diretamente na compressão do vídeo. Experimentos e testes foram feitos utilizando-se o CODEC H.264 desenvolvido no Projeto Plataforma de Convergência Digital IPTV/TV Digital (DigConv) (UNISINOS, 2008a). A qualidade da imagem é medida, através da métrica mais utilizada na literatura o PSNR (Peak Signal to Noise Ratio), que é calculada pelo próprio CODEC ao final da codificação de um vídeo.

Este artigo encontra-se organizado da seguinte forma: na seção 1 é feita a introdução, a seção 2 apresenta a arquitetura do H.264, a seção 3 apresenta a formulação matemática do modelo proposto, a seção 4 descreve o modelo computacional do SMAC, a seção 5 descreve os experimentos e resultados obtidos e por fim na seção 6 são apresentadas as conclusões e sugestões de trabalhos futuros.

## 2. Arquitetura do H.264

A Figura 1 mostra os principais módulos que compõem o padrão de codificação de vídeo H.264/AVC usado no SBTVD.

O padrão de codificação usado no SBTVD (ABNT, 2007), o H.264/MPEG-4 AVC, abrange uma enorme faixa de aplicações que englobam desde vídeos a baixas taxas (vídeo para celulares)

até altas taxas de transmissão (difusão de TV), bem como diversas resoluções espaciais (quantidades de pixels por imagem) e temporais (taxa de quadros ou imagens por segundo).

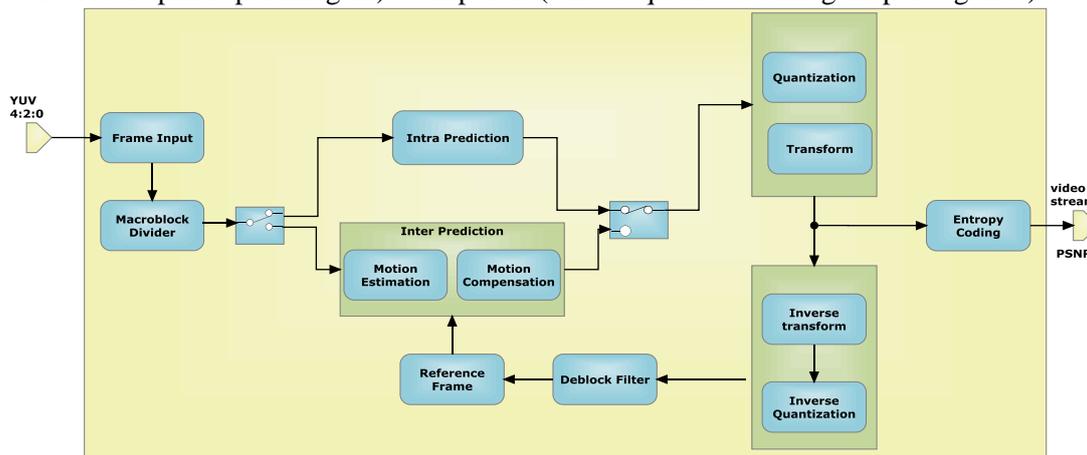


Figura 1: Arquitetura do H.264

O padrão H.264/AVC satisfaz o objetivo de alcançar as mais elevadas taxas de processamento, dentre todos os demais padrões existentes. Mas para isso, foi necessário um grande aumento na complexidade computacional das operações dos CODECs que seguem o padrão H.264/AVC, em relação aos demais padrões disponíveis na atualidade. Este aumento de complexidade impede, na tecnologia atual, a utilização de CODECs H.264/AVC implementados em software quando as resoluções são elevadas e/ou quando se deseja tempo real, com 30 quadros por segundo, por exemplo. A intratabilidade do problema via *software* somado ao enorme interesse comercial que reside neste padrão, têm impulsionado equipes de pesquisa e desenvolvimento ao redor do mundo a tratarem deste tema visando otimizações algorítmicas e/ou implementações em *hardware* para que os requisitos das aplicações sejam atendidos (Silva, 2007).

Considerando a diversidade de conteúdos a serem transmitidos, tais como: apresentação de telejornais, filmes, programas esportivos, etc.; a escolha adequada desses parâmetros se torna ainda mais crítica, pois, devido às suas características, esses conteúdos em geral requerem configurações de codificador e complexidades computacionais bastante distintas para uma mesma qualidade de vídeo pré-definida (IME, UERJ, UFRJ, & UnB, 2010).

A etapa mais custosa do codificador do H.264 é a estimação de movimento, responsável pela busca de descolamentos entre quadros, com um gasto de aproximadamente 90% do tempo total de codificação. Por causa disso, qualquer alteração feita no codificador que implique aumento no tempo gasto para a estimação de movimento terá um impacto final considerável no tempo de execução da codificação. Este é um caso de aumento da complexidade mais perceptível que pode ser provocado pelo aumento no número de partições ou formas de se dividir o macro bloco.

Na Figura 1, temos os módulos que compõem o codificador e decodificador de vídeo H.264. Serão descritos brevemente apenas os módulos do codificador, pois o decodificador não faz parte do escopo deste trabalho. Na arquitetura do codificador temos como entrada um vídeo no padrão YUV 4:2:0 (ITU-T, 2007) que será codificado pelo CODEC H.264/AVC do Sistema Brasileiro de TV Digital (SBTVD). Este vídeo pode ser de três formatos específicos: HD (*High Definition*) or SD (*Standard Definition*) or LD (*Low Definition*). O CODEC H.264 suporta segmentação de figuras no formato de quadros ou *slices* que são formados por macro blocos (ITU-T, 2007; Wenger, 2003). O módulo de partição do macro bloco, representado na Figura 1 pelo Módulo MB, é encarregado de converter os *slices* de vídeo em estruturas de macro blocos (MB) e estes são consideradas as unidades básicas de codificação de vídeo.

O módulo *Intra Prediction* tem por finalidade reduzir a redundância espacial da imagem, onde cada frame é tratado individualmente (ITU-T, 2007).

O módulo *Inter Prediction* reduz a redundância de informações entre quadros denominada redundância temporal. Essa predição consiste em subtrair um quadro de referência (ou quadro de

predição), que seja semelhante ao quadro atual, do próprio quadro atual, para obter uma diferença entre ambos, que será codificada e transmitida (ITU-T, 2007).

As transformadas se baseiam na DCT (*Discrete Cosine Transform*) e têm como objetivo transformar os blocos de imagem (matrizes de *pixels*) em coeficientes. Isto faz com que grande parte da informação do bloco se concentre na área superior esquerda do mesmo, viabilizando o processo de quantização (UNISINOS, 2008a).

A maior parte da distorção é resultante do processo de quantização e pode ser controlada pelo valor do Parâmetro de Quantização.

O processo de quantização consiste basicamente em dividir os coeficientes transformados diminuindo sua energia. Assim, valores menos importantes para a reconstrução da imagem são zerados ou se aproximam de zero (UNISINOS, 2008a).

O processo de codificação por entropia consiste em aplicar os métodos de codificação CABAC (*Context Adaptive Binary Arithmetic Coding*, Codificação Aritmética Binária Adaptável ao Contexto) e CAVLC (*Context-adaptive Variable-Length Coding*, Codificação de Comprimento Variável Adaptável ao Contexto) (UNISINOS, 2008a).

O processo de remontagem de *frame* constrói a *stream* MPEG-4 que será empacotada e enviada pelo meio de transmissão. Como saída desses processos temos o vídeo codificado e o PSNR (*Peak Signal to Noise Ratio*), considerada a métrica objetiva de qualidade de vídeo mais amplamente utilizada (Winkler & Mohandas, 2008).

### 3. Formulação Matemática do Modelo

A Equação (1) representa uma métrica objetiva que avalia a qualidade de vídeo de um CODEC H.264. Quanto maior o valor da FO, maior será a qualidade do vídeo codificado.

As variáveis de decisão correspondem aos parâmetros de configuração do H.264 estudados neste trabalho e estão listadas a seguir:

FR = taxa de *frame rate*;

BR = taxa de *bit rate*;

QI = parâmetro de quantização de *frames* tipo I;

QP = parâmetro de quantização de *frames* tipo P;

QB = parâmetro de quantização de *frames* tipo B;

PF = quantidade de *frames* tipo B em um GOP;

$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$  = pesos pré-definidos e não tendenciosos das variáveis de decisão BR, FR, QI, QP, QB e PF respectivamente;

Segue a formulação matemática do modelo:

#### **Função Objetivo:**

$$\text{Max FO} = \alpha_1 BR + \alpha_2 FR + \alpha_3 \frac{1}{QI} + \alpha_4 \frac{1}{QP} + \alpha_5 \frac{1}{QB} + \alpha_6 \frac{1}{PF} \quad (1)$$

Sujeito a:

$$30 \leq BR \leq 300000 \quad (2)$$

$$15 \leq FR \leq 172 \quad (3)$$

$$0 \leq QI \leq 51 \quad (4)$$

$$0 \leq QP \leq 51 \quad (5)$$

$$0 \leq QB \leq 51 \quad (6)$$

$$0 \leq PF \leq 8 \quad (7)$$

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 > 0 \quad (8)$$

A Função Objetivo, dada pela equação (1), busca a maximização da qualidade de um vídeo codificado no padrão H.262. Esta função é composta por seis variáveis, onde cada uma delas representa um parâmetro de codificação essencial do CODEC H.264/AVC.

A restrição (2) representa os possíveis os valores assumidos pelo *bit rate* ao se codificar um vídeo dentro de um nível/perfil do CODEC H.264. O nível e o perfil do CODEC H.264

estabelecem os valores permitidos para alguns parâmetros de codificação dentre eles o *bit rate*. Os valores de *bit rates* permitidos pelo CODEC H.264 e que estão de acordo com seu nível e perfil são mostrados na Tabela 1.

Tabela 1 – Variação do bit rate de acordo com o nível e perfil do H.264.

Nível (Profile Level)	Bit rate (Kbps)	
	Perfil	
	Baseline, Main	High
1	30 a 64	60 a 80
1b	64 a 128	80 a 160
1.1	128 a 192	160 a 240
1.2	192 a 384	240 a 480
1.3	384 a 768	480 a 960
2	768 a 2000	960 a 2500
2.1	2000 a 4000	2500 a 5000
2.2	2000 a 4000	2500 a 5000
3	4000 a 10000	5000 a 12500
3.1	10000 a 14000	12500 a 17500
3.2	14000 a 20000	17500 a 25000
4	14000 a 20000	17500 a 25000
4.1	20000 a 50000	25000 a 62500
4.2	20000 a 50000	25000 a 62500
5	50000 a 135000	62500 a 168750
5.1	135000 a 240000	168750 a 300000

Fonte: ITU-T (2007).

As restrição (3) representa os possíveis valores assumidos pelo *frame rate* ao se codificar um vídeo dentro de um determinado formato e nível do CODEC H.264. Esses valores são mostrados na Tabela 2.

Tabela 2 – máximo *frame rate* de acordo com o nível do H.264.

Nível (Profile Level)	QCIF 176x144	CIF 352x288	525SD 720x480	720pHD 1280x720	1080HD 1920x1088
1	15	-	-	-	-
1b	15	-	-	-	-
1.1	30,3	7,6	-	-	-
1.2	60,6	15,2	-	-	-
1.3	120	30	-	-	-
2	120	30	-	-	-
2.1	172	50	-	-	-
2.2	172	51,1	15	-	-
3	172	102,3	30	-	-
3.1	172	172	80	30	-
3.2	172	172	161	60	-
4	172	172	172	68,3	30
4.1	172	172	172	68,3	30
4.2	172	172	172	145,1	63,8
5	172	172	172	163,8	72,3
5.1	172	172	172	172	120,5

Fonte: ITU-T (2007).

As restrições (4), (5) e (6) garantem que os parâmetros QI, QP e QB possuam valores dentro do intervalo de números inteiros positivos [0, 51]. A restrição (7) garante que a quantidade de frames tipo B dentro de um GOP (Group Of Picture) de tamanho igual a 10 frames, varie dentro de um intervalo de números inteiros positivos [0, 8].

A restrição (8) garante que os pesos das variáveis de decisão sejam maiores que zero. Esses pesos foram previamente calculados a fim de tornar a equação (1), uma solução não tendenciosa (SNT). Para calcular os pesos da SNT, o aplicativo SMAC foi executado 100 vezes onde foram atribuídos valores aleatórios inteiros de 0 a 100 aos pesos, seguindo uma distribuição normal. A BT usou a seguinte configuração: nbmax igual a 100 (critério de parada) e lista tabu igual a 7. O AG utilizou uma taxa de crossover igual a 0,8 e taxa de mutação igual a 0,2. A mesma solução inicial foi usada em todas as execuções. Após 100 rodadas do SMAC, tirou-se uma média dos valores assumidos pelas variáveis de decisão, onde a variável de maior média teve seu peso normalizado no valor 1 e o peso das demais variáveis foi calculado dividindo-se a maior média encontrada pelas médias de cada variável de decisão, sendo o resultado desta divisão então arredondado e atribuído como peso não tendencioso a cada uma das respectivas variáveis de decisão da FO. Os pesos não tendenciosos encontrados foram:  $\alpha_1 = 7$ ,  $\alpha_2 = 1$ ,  $\alpha_3 = 32$ ,  $\alpha_4 = 24$ ,  $\alpha_5 = 21$ ,  $\alpha_6 = 122$ , que correspondem respectivamente aos pesos do bit rate, frame rate, quantização de quadros I, quantização de quadros P, quantização de quadros B, quantidade de quadros B em um GOP.

#### 4. Modelo Computacional do SMAC

Uma arquitetura híbrida é uma combinação de técnicas usada para se obter melhores resultados. Neste caso serão usadas as metaheurísticas Busca Tabu (BT) e Algoritmo Genético (AG) para compor a arquitetura do modelo computacional do SMAC que é mostrada na figura 2. As pesquisas estão se voltando ao uso de técnicas híbridas (Burke, De Causmaecker, & Vanden Berghe, 1999), pois elas têm apresentado melhores resultados, visto que exploram as vantagens de cada método aplicado. Na literatura foram encontrados estudos (Xu, Bi, & Mao, 2000; Yasakethu, et al., 2008), onde se aplica o Algoritmo Genético em módulos específicos de um CODEC H.264. Yuelel et al (2000) aplica um AG no módulo de Estimção de Movimento do CODEC enquanto que em Yasakethu et al (2008) encontramos um AG aplicado ao módulo de Quantização. Em Koumaras et al (2005) foi estudado o comportamento do *bit rate* e *frame rate* para vídeos do tipo CIF e QCIF a fim de criar uma métrica objetiva para avaliar a qualidade de vídeo percebida.

O modelo computacional proposto na Figura 2 tem como finalidade encontrar dentro de um espaço de busca, uma melhor configuração de parâmetros para o CODEC H.264 e que vão atuar sobre a arquitetura sistêmica mostrada na Figura 1. A ideia principal é que as duas metaheurísticas, BT e AG trabalhem em conjunto para encontrar uma boa para o problema de configuração de parâmetros do CODEC H.264/AVC do SBTVD.

A arquitetura híbrida do SMAC é composta por 4 módulos: módulo que recebe a solução inicial e configura o modelo para trabalhar em um nível e perfil específico do padrão H.264 para um formato de vídeo especificado; módulo da Busca Tabu, módulo do Algoritmo genético e o módulo do CODEC H.264.

Primeiramente é fornecido para o modelo uma solução inicial viável e um vídeo com determinado número de frames no formato YUV 2:2:0. Esta solução é composta pelos parâmetros de configuração do H.264 estudados, que são: *bit rate*; *frame rate*; parâmetros de quantização para quadros tipo B, P e I; e quantidade de slices tipo B dentro em um GOP (*Group of Pictures*).

O módulo da Busca Tabu vai operar sobre o conjunto de parâmetros do CODEC H.264 que fazem parte da solução inicial a fim de obter um conjunto das melhores soluções viáveis, chamadas soluções de elite e correspondem às soluções que apresentaram a melhor função objetivo. A BT utiliza como função objetivo a equação (1). A quantidade de soluções de elite geradas pela BT é de no máximo 20 soluções. Caso a BT atinja o valor do nbmax sem, contudo ter gerado 20 soluções então essas soluções são repassadas em menor quantidade ao Algoritmo Genético que vai dar a elas um tratamento específico. O módulo de geração de vizinhança da BT vai gerar 50 vizinhos a cada iteração. A BT utiliza uma lista tabu composta por no máximo sete movimentos tabu. Essa lista guarda uma solução completa, ou seja, os valores de todos os

parâmetros que fazem parte de uma mesma solução, fazendo com que essa solução seja considerada um movimento tabu.

O módulo de critério de aspiração da BT prevê que será aceito um movimento, mesmo que tabu, se ele melhorar o valor da função objetivo global. O critério de parada utilizado pela BT é o nbmax de tamanho igual a 100. Quando a BT atinge seu critério de parada então ela passa o controle para o Algoritmo Genético (AG) e envia a ele o conjunto de soluções de elite.

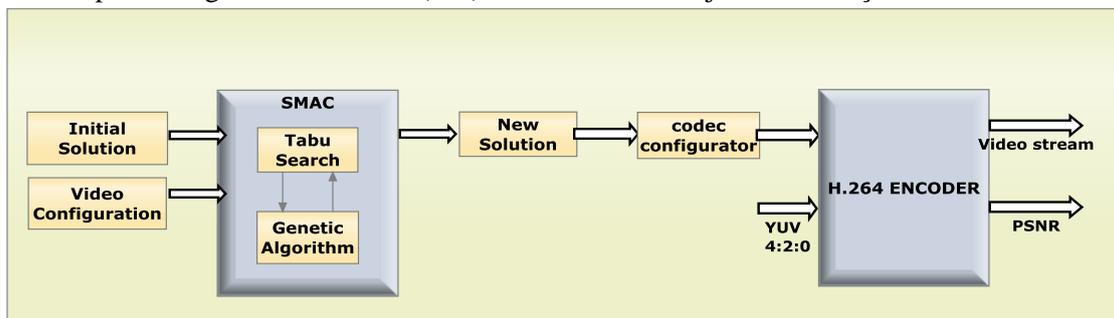


Figura 2: Modelo computacional do algoritmo SMAC.

O módulo Algoritmo genético (AG) recebe como entrada o conjunto de soluções de elite geradas pela BT para ser sua população inicial. Caso a BT tenha enviado ao AG uma quantidade de soluções de elite menor que 20, então o módulo de geração de vizinhança do AG complementa sua população inicial gerando aleatoriamente quantos indivíduos forem necessários para se chegar a uma população inicial de 20 indivíduos. O AG utiliza como função de fitness a equação (1). Esta função avalia o grau de aptidão de um indivíduo da população, neste caso, quanto maior a *fitness*, melhor a qualidade da solução encontrada. O cromossomo é representado pelo conjunto dos seis parâmetros que compõem a solução inicial  $S_0$ . O AG utiliza a seleção por torneio, além de um cruzamento aritmético com taxa de 0,8 e a mutação uniforme com taxa de 0,2. O critério de parada do AG é o número de gerações igual a 100.

Quando o AG atinge seu número de gerações o modelo SMAC compara se houve uma melhora na função *fitness* em relação a melhor função objetivo da BT, caso tenha havido essa melhora então o SMAC aciona novamente a BT passando para ela como solução inicial a melhor solução encontrada pelo AG. Caso não tenha ocorrido uma melhora na melhor solução encontrada até o momento, o SMAC termina a execução do algoritmo e torna a função objetivo da BT a melhor solução encontrada pelo algoritmo a qual será usada nas configurações do CODEC H.264.

## 5. Experimentos e resultados

Na primeira etapa dos experimentos foi executado o algoritmo SMAC para cada um dos vídeos apresentados na Tabela 3. As melhores soluções encontradas pelo SMAC foram usadas para configurar o CODEC H.264 do projeto DigConv (UNISINOS, 2008a, 2008b). Os vídeos foram codificados pelo CODEC H.264 previamente configurado e como resultado desta codificação obteve-se o PSNR e o grau de compressão desses vídeos. Esses resultados foram analisados e comparados com os resultados obtidos pela solução inicial.

Tabela 3 – Vídeos usados nos experimentos.

Video	Formato	Resolução	Perfil / Nível	Nro Frames	Tamanho Original (KB)
V02	720p5994 HD - Parkrun	1280 x 720	High/ 40	253	342.688
V10	CIF - Bus	352 x 288	High/ 40	150	22.275
V15	CIF - News	352 x 288	High/ 12	300	44.550
V29	720p HD - Parkrun	1280 x 720	High/ 40	10	342.784
V33	CIF - Bridge close	352 x 288	Main/ 20	50	297.008
V34	SD525 - Pedestrian	720 x 480	Main/ 30	100	50.625

A Tabela 3 relaciona os tipos de vídeos utilizados no experimento. Na primeira coluna tem-se a identificação do vídeo, ou seja, como ele será referenciado daqui por diante; na segunda coluna tem-se o formato e o nome do vídeo, onde o formato se refere ao formato final que esse vídeo será codificado pelo CODEC H.264, lembrando que todo vídeo que entra no CODEC deve estar obrigatoriamente no formato YUV 4:2:0; na terceira coluna tem-se a resolução do vídeo; na quarta coluna temos o perfil e o nível do H.264 nos quais o vídeo foi codificado, na quinta coluna tem-se a quantidade de frames codificados e na sexta coluna tem-se o tamanho original do vídeo dado em Kbytes quando ele está no seu formato original, neste caso, o formato YUV 4:2:0.

Em resumo, um mesmo vídeo foi codificado pelo CODEC H.264 em dois momentos distintos, no primeiro momento a equipe DigConv codificou o vídeo usando valores padrões para o conjunto de parâmetros de configuração do CODEC H.264. Estes parâmetros são os que fazem parte da Solução Inicial. No segundo momento, o vídeo foi codificado usando o conjunto de parâmetros obtidos pela melhor FO encontrada pelo SMAC. O PSNR e o grau de compressão obtidos a partir dessas duas codificações foram então comparados e analisados e os resultados são apresentados na Tabela 4. Esta mostra os resultados dos experimentos obtidos após a codificação dos vídeos que estão relacionados na Tabela 3 e que foram codificados pelo CODEC H.264. Cada linha da tabela 4 corresponde aos valores dos parâmetros de configuração do CODEC H.264 que foram estudados (colunas 3 a 8 da Tabela 4), bem como os resultados obtidos em termos de PSNR e grau de compressão desse vídeo (colunas 13 e 14). A primeira solução, apresentada para cada vídeo da tabela 4 (coluna 1) corresponde à Solução Inicial proveniente da codificação desse vídeo pela equipe DigConv. Os resultados obtidos pela solução inicial, que são o PSNR e o grau de compressão do vídeo, são comparados aos resultados das soluções encontradas pelo SMAC. As demais linhas da tabela, que estão agrupadas em um determinado vídeo, correspondem aos parâmetros de configuração do CODEC e os resultados da codificação do vídeo obtidos através da melhor solução encontrada pelo SMAC.

Tabela 4 – Vídeos codificados no padrão H.264.

Vídeo	nbmax / Lista Tabu	BR (Kbps)	FR (fps)	QI	QP	QB	PF	Média PSNR 1	Média PSNR N	FO	Tamanho vídeo (KB)	% PSNR N	% compressão
V02	Solução Ini.	54.400,25	30,00	28	28	30	1	43,383	37,042	380.956,4500	19.694		
	100 / 50	62.499,96	67,46	30	28	30	1	43,319	36,937	437.691,8095	19.511	-0,28%	0,93%
	400 / 100	62.499,99	68,06	27	26	28	1	43,470	37,790	437.692,8355	26.465	2,02%	-34,38%
	600 / 100	62.499,98	67,89	28	30	35	1	43,380	36,340	437.692,3239	12.951	-1,90%	34,24%
V10	Solução Ini.	1.600,35	30,00	28	28	0	0	39,109	39,589	11.234,4500	977		
	50 / 7	19.999,7900	99,12	26	26	21	1	40,298	41,484	140.222,8071	1.764	4,79%	-80,55%
	200 / 100	62.488,3203	99,12	26	26	21	1	40,298	41,484	437.642,5193	1.542	4,79%	-57,83%
	400 / 100	62.499,9922	99,87	26	25	21	1	40,298	41,097	437.725,0055	793	3,81%	18,83%
	800 / 100	62.499,9883	99,45	27	25	29	1	40,298	40,546	437.724,2344	937	2,42%	4,09%
V15	Solução Ini.	246,31	15,00	28	28	30	1	37,357	43,571	1.863,8700	600		
	100 / 60	479,9928	15,16	26	25	27	1	37,357	43,480	3.500,077	594	-0,21%	1,00%
	100 / 50	479,9954	15,06	28	28	30	1	37,357	43,554	3.499,732	598	-0,04%	0,33%
V29	Solução Ini.	25.469,40	60,00	28	28	30	2	43,439	35,183	178.409,500	1.451		
	100 / 50	62.499,9375	67,93	29	28	30	1	43,439	35,222	437.692,150	869	0,11%	40,11%
V33	Solução Ini.	1.995,62	30,00	28	28	30	1	43,036	39,763	14.124,040	804		
	100 / 50	3.999,9575	29,92	29	29	30	1	39,639	39,037	28.154,254	607	-1,83%	24,50%
V34	Solução Ini.	10.036,89	30,00	28	28	30	1	43,101	45,601	70.412,9300	4.044		
	100 / 50	13.999,9141	29,98	29	30	31	1	43,101	45,609	98.153,960	4.057	0,02%	-0,32%

Em alguns casos foram testadas mais de uma solução do SMAC para um mesmo vídeo. Todas as soluções encontradas pelo SMAC foram obtidas, após 100 rodadas completas do algoritmo. Para cada rodada completa do SMAC, foram configurados o nbmax e a Lista Tabu, apresentadas na coluna 2 da tabela, de forma distinta, para se avaliar as diversas soluções encontradas e tentar alcançar uma melhor solução para a parametrização do CODEC e consequentemente uma melhor codificação do vídeo. Na coluna 9 temos a média do PSNR correspondente ao primeiro frame codificado e na coluna 10 temos a média do PSNR de todos os

frames codificados. A coluna 11 apresenta o resultado da Função Objetivo e na coluna 12 temos o tamanho final do vídeo após sua codificação pelo H.264. O tamanho original dos vídeos encontra-se na Tabela 3.

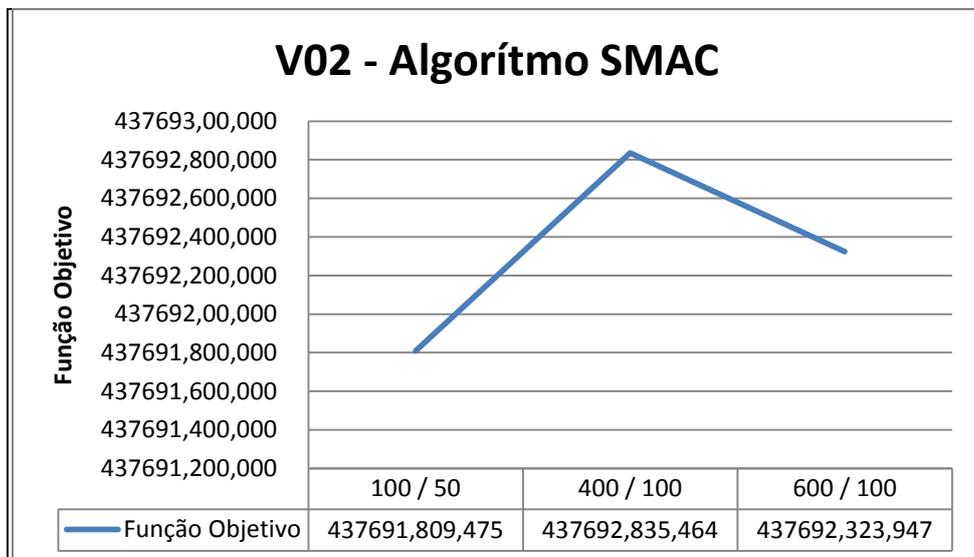


Figura 3 – Algoritmo SMAC aplicado ao Vídeo V02.

A Figura 3 mostra a evolução dos valores das Funções Objetivo que foram obtidas para o vídeo V02 (Tabela 4) usando-se para isso, diferentes configurações de valores de nbmax e tamanho da Lista Tabu para o algoritmo híbrido SMAC. Observou-se que o melhor valor da FO foi alcançado quando o SMAC utilizou um nbmax igual a 400 e um tamanho de Lista Tabu igual a 100.

A Figura 4 mostra os resultados alcançados em termos de PSNR e de compressão de vídeo para o vídeo V02 (conforme Tabela ). Esses resultados foram obtidos quando se configurou o CODEC H.264 de acordo com as soluções encontradas pelo SMAC. Observando-se o gráfico da Figura 4 temos na linha horizontal as soluções 1, 2 e 3 que correspondem às FOs apresentadas no gráfico da Figura 3.

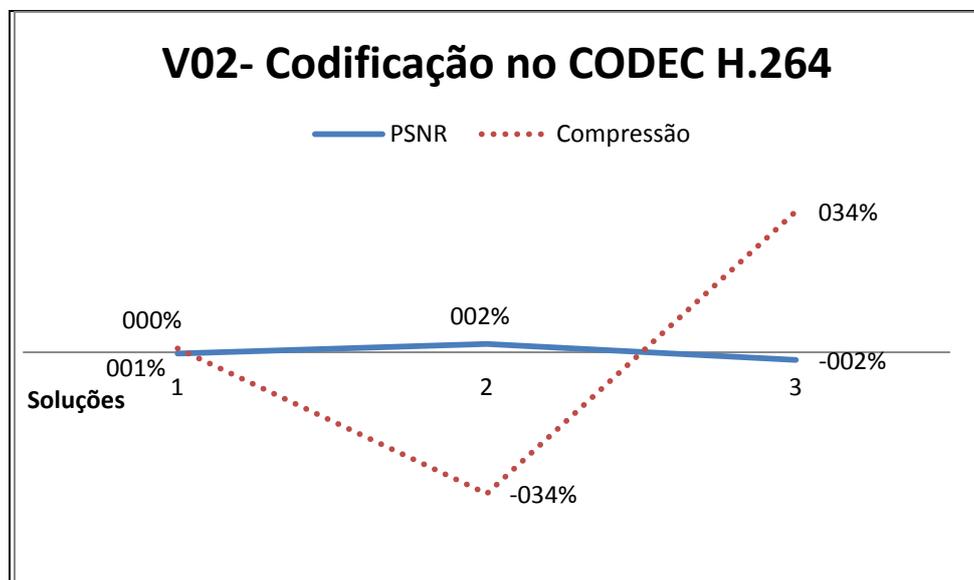


Figura 4 – Resultados da codificação do vídeo V02.

Sobrepondo-se os gráficos das figuras 3 e 4 observa-se que a melhor FO corresponde à da solução 2. Na solução 2, encontramos o mais alto ganho de PSNR que foi de 2,02% e uma perda de compressão significativa, de 34,38%. Isso ocorreu porque o SMAC prioriza o ganho de qualidade de imagem que se reflete no aumento do PSNR.

De acordo com a Tabela 4, observou-se que os vídeos V02 e V10 obtiveram seus melhores valores da FO quando o SMAC usou um nbmax igual a 400 e tamanho da Lista Tabu igual 100, o que leva a crer que essa é uma boa configuração para a Busca Tabu.

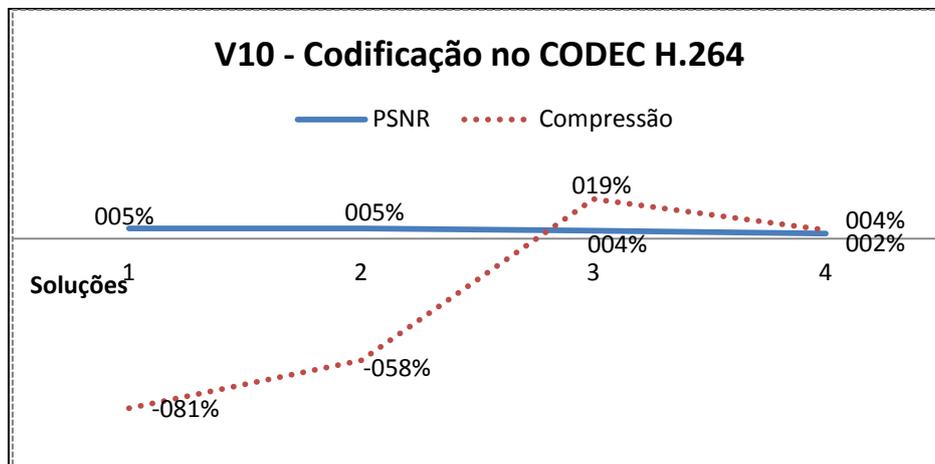


Figura 5 – Resultados da codificação do vídeo V10.

A Figura 5 mostra os resultados alcançados em termos de PSNR e de compressão para o vídeo V10 conforme Tabela .

Na Figura 5 fica claro que a dinâmica do SMAC busca dar prioridade ao ganho de PSNR e em seguida ao ganho de compressão. Nota-se que as soluções 1 e 2 para o vídeo V10 obtiveram o mesmo ganho de PSNR, porém para o SMAC o melhor valor da FO entre as duas soluções foi aquela em que se conseguiu minimizar a perda de compressão, que é o caso da solução 2. Outro comportamento interessante, que pode ser observado na Figura 5, é que as soluções 3 e 4 alcançaram ganhos de PSNR e de compressão, porém a melhor FO foi aquela que obteve o maiores ganhos em ambos, que é o caso da solução 3, que obteve um ganho de compressão em 18,83% e um aumento de 3,81% no PSNR.

O vídeo V15, conforme apresentado na Tabela 4, primeiramente obteve uma perda de 0,21% no PSNR e um ganho de compressão de 1%. Observou-se com isso que o SMAC apesar de não conseguir um ganho no PSNR, alcançou um maior grau de compressão do vídeo e tentou minimizar ao máximo a perda do PSNR. Na segunda solução encontrada para V15, houve uma perda de 0,04% no valor do PSNR, que foi uma perda menor em relação à primeira solução, e um ganho de 0,33% de compressão. Observou-se que a primeira solução obteve um melhor valor da FO em relação à FO da segunda solução. Analisando-se essas duas FOs fica claro que quando o SMAC não alcança um ganho no PSNR, ele então tenta maximizar a compressão. Esse comportamento é o mesmo encontrado quando se analisa a primeira e segunda solução do vídeo V02, ou seja, as duas FOs obtiveram perda de PSNR, porém a melhor solução foi aquela obteve maior compressão de vídeo.

De acordo com a Tabela , os vídeos V29, V33 e V34 foram codificados usando-se soluções que foram encontradas pelo SMAC e este utilizou nbmax igual 100 e tamanho da Lista Tabu igual 50. O resultado da codificação do V29 obteve ganhos de 0,11% no PSNR e de 40,11% na compressão, o que é um ótimo resultado, pois não houve perdas, nem de qualidade de imagem nem de compressão. O resultado da codificação do V33 obteve uma perda de 1,83% no PSNR mas em compensação alcançou um ganho de compressão de 24,50%. O resultado da codificação do V34 obteve um pequeno ganho de 0,02% no PSNR e uma pequena perda de compressão de 0,32%.

Constatou-se em todos os experimentos que o comportamento do SMAC independe do formato do vídeo. Os maiores ganhos de compressão foram obtidos quando se codificou um vídeo no formato HD, que é o caso dos vídeos V02 e V29 mostrados na Tabela . As soluções encontradas, para os vídeos V02 e V29, pelo SMAC conseguiram um ganho de 34,24% e 40,11% respectivamente no grau de compressão desses vídeos. O tempo médio de uma execução completa do SMAC foi de 4 minutos, sabendo-se que uma rodada completa equivale a 100 execuções do SMAC.

## 6. Conclusão

Foi apresentado um novo algoritmo híbrido denominado Simulador de Metaheurísticas aplicado a um CODEC (SMAC) que utiliza a Busca Tabu em combinação com o Algoritmo Genético com o objetivo de estudar o comportamento de seis dos principais parâmetros de configuração do CODEC de vídeo H.264/AVC. Os parâmetros estudados foram: *bitrate*; *frame rate*; parâmetros de quantização para quadros tipo B, P e I; e quantidade de quadros tipo B dentro em um GOP. O algoritmo proposto busca a melhor configuração desses parâmetros no intuito de poder configurar o CODEC H.264/AVC do Sistema de Televisão digital Brasileiro e assim alcançar uma melhora na qualidade da imagem de um vídeo codificado neste padrão. O modelo SMAC provou ser robusto no sentido de poder pesquisar a melhor configuração do CODEC H.264 para vários tipos de vídeo codificados nos diferentes perfis do padrão H.264 em todos os seus níveis. O modelo foi validado através dos padrões de comportamento dos parâmetros encontrados na literatura. A dinâmica de comportamento do SMAC traduz a dinâmica de CODEC real (JIGA de testes). As variáveis de decisão estudadas se mostraram bastante representativas, pois viabilizaram uma correta análise em termos de qualidade de imagem e compressão de vídeo. É nitidamente observado o ganho computacional que esta ferramenta proporciona ao CODEC H.264, pelo fato dela ter o poder de melhorá-lo sem que seja necessária a alteração de seu código-fonte. O simulador proposto consegue maximizar a performance do referido CODEC trabalhando apenas com a configuração adequada de seus parâmetros mais representativos.

Como trabalhos futuros sugere-se aprimorar o modelo acrescentando outras variáveis de decisões que influenciem diretamente na qualidade da imagem, como por exemplo, a quantidade de partições de um macro bloco, a dinâmica de um vídeo, a quantidade de frames de referência a serem considerados. Esses são apenas alguns dos inúmeros parâmetros de configuração do H.264 que podem ser estudados e acrescentados neste trabalho.

## Referências

- ABNT, NBR 15602-1. (2007). Televisão Digital Terrestre – Codificação de vídeo, áudio e multiplexação, Parte 1: Codificação de vídeo. In Rio de Janeiro: ABNT.
- Burke, E., De Causmaecker, P., & Vanden Berghe, G. (1999). A Hybrid Tabu Search Algorithm for the Nurse Rostering Problem. In B. McKay, X. Yao, C. Newton, J.-H. Kim & T. Furuhashi (Eds.), (Vol. 1585, pp. 187-194): Springer Berlin / Heidelberg.
- Cermak, G., Pinson, M., & Wolf, S. (2011). The Relationship Among Video Quality, Screen Resolution, and Bit Rate. *Broadcasting, IEEE Transactions on*, 57, 258-262.
- Correia, P. L., & Pereira, F. (2003). Objective evaluation of video segmentation quality. *Image Processing, IEEE Transactions on*, 12, 186-200.
- CPqD. (2006a). Arquitetura de Referência. Sistema Brasileiro de Televisão Digital Terrestre. In PD.30.12.34A.0001A/RT-13/AA (Ed.).
- CPqD. (2006b). Especificação Técnica de Referência. In 30.12.34A.0001A/RT-14/AA (Ed.).
- Fiems, D., Steyaert, B., & Bruneel, H. (2012). A genetic approach to Markovian characterisation of H.264 scalable video. *Multimedia Tools and Applications*, 58, 125-146.
- IME, UERJ, UFRJ, & UnB. (2010). Projeto H.264 – SETUP. Linhas Mestras para Operação e Configuração de Sistemas Compressão de Vídeo para o SBTVD. In.
- ITU-R. (2002). Recommendation ITU-R BT.500-11: Methodology for the subjective assessment of the

- quality of television pictures. In: Swiss.
- ITU-T. (2007). ITU-T Recommendation H.264. Advanced video coding for generic audiovisual services. In: ITU-T.
- Jianning, Z., Yuwen, H., Shiqiang, Y., & Yuzhuo, Z. (2003). Performance and complexity joint optimization for H.264 video coding. In *Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on* (Vol. 2, pp. II-888-II-891 vol.882).
- Malvar, H. S., Hallapuro, A., Karczewicz, M., & Kerofsky, L. (2003). Low-complexity transform and quantization in H.264/AVC. *Circuits and Systems for Video Technology, IEEE Transactions on, 13*, 598-603.
- Moriyoshi, T., Shinohara, H., Miyazaki, T., & Kuroda, I. (2001). Real-Time Software Video Codec with a Fast Adaptive Motion Vector Search. *The Journal of VLSI Signal Processing, 29*, 239-245.
- Nemethova, O., Ries, M., & Rupp, M. (2004). Quality Assessment for H.264 Coded Low-Rate and Low-Resolution Video Sequences. In *Vortrag: IASTED International Association of Science and Technology for Development* (pp. 136-140). St. Thomas, US Virgin Islands: Proceedings of 3rd Conference on Communications, Internet, and Information Technology (CIIT 2004).
- Sikora, T. (2005). Trends and Perspectives in Image and Video Coding. *Proceedings of the IEEE, 93*, 6-17.
- Silva, A. M. c. d. (2007). *Um estudo sobre o padrão H.264/AVC de compressão de vídeo*. Universidade Católica de Pelotas.
- Tianbing, X., & Weidong, C. (2006). A Fast Adaptive Statistical Genetic Motion Search Algorithm for H.264/AVC^1. In *Advanced Information Networking and Applications, 2006. AINA 2006. 20th International Conference on* (Vol. 1, pp. 553-558).
- UNISINOS. (2008a). Especificação de Software CODEC – Codificação e Decodificação de Sinais Fonte. In: UNISINOS - Universidade do Vale do Rio dos Sinos.
- UNISINOS. (2008b). Especificação Técnica de Hardware – Codificação e Decodificação de Sinais Fonte. In: UNISINOS - Universidade do Vale do rio dos Sinos.
- Wenger, S. (2003). H.264/AVC over IP. *Circuits and Systems for Video Technology, IEEE Transactions on, 13*, 645-656.
- Winkler, S., & Mohandas, P. (2008). The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics. *Broadcasting, IEEE Transactions on, 54*, 660-668.
- Xu, Y., Bi, D., & Mao, B. (2000). A genetic search algorithm for motion estimation. In *Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on* (Vol. 2, pp. 1058-1061 vol.1052).
- Yasakethu, S. L. P., Fernando, W. A. C., Adedoyin, S., & Kondoz, A. (2008). A rate control technique for off line H.264/AVC video coding using subjective quality of video. *Consumer Electronics, IEEE Transactions on, 54*, 1465-1472.
- Yu-Wen, H., Bing-Yu, H., Shao-Yi, C., Shyh-Yih, M., & Liang-Gee, C. (2006). Analysis and complexity reduction of multiple reference frames motion estimation in H.264/AVC. *Circuits and Systems for Video Technology, IEEE Transactions on, 16*, 507-522.