

## AVALIAÇÃO DE ALGORITMOS PARA A ESTIMAÇÃO DE MOVIMENTO APLICADOS NA PRÓXIMA GERAÇÃO DE VÍDEOS EM ULTRA-ALTA DEFINIÇÃO

CÁSSIO CRISTANI<sup>1,2</sup>; PARGLES DALL'OGGIO<sup>1</sup>; BYEUNGWOO JEON<sup>2</sup>;  
MARCELO PORTO<sup>1</sup>, LUCIANO AGOSTINI<sup>1</sup>

<sup>1</sup>Universidade Federal de Pelotas – {cristani, pdalloggio, mporto, agostini} @inf.ufpel.edu.br

<sup>2</sup>Sungkyunkwan University – bbjeon@skku.edu

### 1. INTRODUÇÃO

A América Latina tem uma fatia de 7,8% do mercado mundial de televisores, que comercializa mundialmente cerca de R\$ 850 bilhões. O Brasil vem expandindo-se rapidamente e sozinho comporta uma fatia de 3,43% do mercado mundial, sendo um dos grandes compradores mundiais de televisores (IDATE, 2011). Nesse cenário é de extrema importância a criação de tecnologia em território nacional, visto que o produto tem um preço mais competitivo, mais empregos sendo ofertado, profissionalização dos serviços e o capital entra em circulação fazendo a indústria servir de suporte a economia nacional.

Os vídeos em altas definições (HD), atualmente disseminados em diversos dispositivos como celulares, monitores, painéis e na própria internet foram inseridos no mercado por meio de televisores e popularizados pela qualidade que trazem ao usuário. Como qualquer vídeo digital, estes são imagens estáticas representadas a uma taxa de 24 a 30 vezes por segundo possibilitando assim que o Sistema Visual Humano note um movimento suave, ou seja, sem cortes ou pausas. Contudo, a qualidade superior que pode representar detalhes imperceptíveis anteriormente traz o ônus de aumentar largamente o tamanho final do vídeo, fazendo com que estes vídeos não sejam possíveis de ser transmitidos e processados em sua forma pura. Mesmo os transponders de satélites mais potentes do mundo - que tem uma capacidade em torno de 400 Mbps - não teriam capacidade de transmitir em tempo real. Para solucionar esse gargalo os vídeos passam por um processo de codificação, ou compactação, para torna-los viáveis de serem manipulados.

Se analisarmos pela ótica da rede de internet, temos uma figura clara da importância dos vídeos atualmente. O tráfego de vídeos estará consumindo 69% da capacidade mundial no ano de 2017. No ano de 2012, os vídeos usavam 57% de banda entre todas as transferências realizadas. Contudo, esse dado não contabiliza as transferências Peer-to-peer (P2P), que são feitas diretamente de usuário para usuário e tem um grande impacto no fluxo total de dados. As estimativas de (CISCO, 2013) preveem que em quatro anos os vídeos ocuparão de 80 a 90% da banda de internet considerando P2P, já que o volume tem crescido por conta da resolução, pelo número de usuários, aplicações e dispositivos. Nesse cenário, a compactação destes vídeos se torna indispensável, sem ela teríamos que trocar constantemente a infra-estrutura de rede que além de custosa, em termos financeiros e práticos, tem um limite teórico que impediria a melhoria na qualidade dos vídeos com o mesmo ritmo que é praticado atualmente.

### 2. METODOLOGIA

Com o aumento da resolução o tamanho do vídeo teve cresceu drasticamente e o codificador H.264 (KUHN, 1999) se tornou obsoleto, pois não conseguia tratar de forma eficiente vídeos em qualidade Full HD, por exemplo. Com a pressão mercadológica para o suporte a Ultra-Alta Definição (UHD), o

vídeo se tornou quatro ou dezesseis vezes maior que um Full HD e o H.264 já não teria condições de codificar estas definições. Com isso, se fez necessário a criação de um novo padrão com taxas de compressão mais robustas, capazes de comprimir com eficiência vídeos em qualidade HD e UHD. O JVC e ITU se responsabilizaram por criar o HEVC (SULLIVAN, 2012), atual estado-da-arte em codificação de vídeo, que propunha obter taxas de compressão duas vezes melhores em relação ao H.264, mantendo a mesma qualidade visual do vídeo.

No Grupo de Arquiteturas e Circuitos Integrados (GACI) da UFPEL, pesquisávamos a etapa mais importante da codificação, que é a estimação de movimento (EM), responsável por reduzir as redundâncias temporais presente em vídeos digitais. Estas análises eram feitas de forma isolada e independente de padrão, o que possibilita a migração destes algoritmos para qualquer padrão existente e mesmo para padrões futuros. A EM é o processo mais complexo de um codificador, representando cerca de 80% dos recursos. Entretanto, nela são obtidos os melhores resultados. Assim, quando produzimos estratégias que melhorem este módulo, o impacto é magnificado por se tratar do caminho crítico do codificador.

Nossas premissas básicas advêm de trabalhos prévios com base em vídeos HD e Full HD, que dão uma confiabilidade grande já que a dificuldade de codificação nestas resoluções é aumentada pelo fenômeno dos mínimos locais, cujos quais são desprezados pela maioria dos autores da área que avaliam os algoritmos com base em vídeos de baixa definição, o que não garante a mesma eficiência para vídeos em resoluções maiores. Entretanto, até este momento todos os resultados gerados eram com base em um módulo isolado, e poderiam sofrer alterações quando integrado ao fluxo completo de codificação. Contudo, os resultados atuais vão de encontro com nossas conclusões previas (PORTO, 2012) (CRISTANI, 2012).

Para validar nossas investigações nos engajamos com a emergência do padrão HEVC e a mobilização por conta de órgãos internacionais em torno da nova tecnologia de Ultra-Alta Definições (UHD 4K e 8K) e inserimos nossos algoritmos no software de referência do HEVC visando extrair o comportamento de nossas estratégias juntamente com as técnicas estado-da-arte, que viabilizaram a melhoria da compressão em duas vezes relativa ao codificador H.264/AVC para uma mesma qualidade visual. Para isso a complexidade computacional aumentou e o paralelismo é fortemente recomendado para não tornar o processo linearmente mais lento. Nessa nova etapa de projeto estamos começando a investigação de vídeos em UHD, que trazem novas características e uma dificuldade muito maior por ser tecnologia para as próximas gerações de TV digital e o hardware atual ter limites práticos para tratar esses problemas desta magnitude.

### **3. RESULTADOS E DISCUSSÃO**

Este trabalho traz a tona os primeiros resultados de uma profunda análise que está começando a ser realizado com vídeos em qualidade UHD, tema que mundialmente vem sendo muito investigado na área de comunicação via satélites, visto a quantidade de dados nunca experimentada até então. Contudo, vem sendo pouco explorado em aplicações específicas como é o caso da Estimação de Movimento que não possui nenhum trabalho de referência para esta nova tecnologia. Isso ocorre principalmente porque o algoritmo rápido de referência (EPZS) vem sendo amplamente utilizado já que chega a resultados muito próximos da solução ótima. Assim, há poucos pesquisadores que ainda estão explorando melhorias com novas estratégias de busca para comparar com o EPZS.

Como nossa preocupação é gerar soluções eficientes em hardware e com alto potencial de paralelismo, continuamos explorando a área propondo soluções que consigam se aproximar em qualidade, sem extrapolar o volume de dados e oferecendo uma complexidade muito menor, com alto grau de paralelismo, que também pode ser explorado em Software, e que sejam regulares viabilizando projetos arquiteturais que buscam alta eficiência energética e aplicações em tempo real.

Neste trabalho fizemos uma análise comparativa entre algoritmos rápidos para a Estimção de Movimento no Software de referencia do padrão HEVC. Entre os algoritmos analisados estão o EPZS, referencia da EM no padrão; o DS e FSS, que são algoritmos clássicos bem conhecidos da comunidade científica, e o *Multiple Iterative Randon Search* (Dall'Oglio, 2012), que é um algoritmo produzido no grupo com o melhor custo-benefício entre qualidade e esforço computacional focado para resoluções até Full HD.

Para os resultados apresentados na Tabela 1, as simulações utilizaram uma área de busca de 64x64 pixels e uma sequencia progressiva de 30 quadros para dois vídeos em Full HD, um vídeo WQXGA e um vídeo UHD 4K. Cada algoritmo tem um resultado de PSNR que é a qualidade objetiva medida em decibéis, e de Bitrate que é a quantidade total de bits gerado após a codificação, que representa o volume total de dados.

**Tabela 1. Comparação algorítmica em qualidade e volume de dados para diferentes resoluções**

Resolução	Full HD				WQXGA		UHD 4K	
Vídeo	Kimono		BQTerrace		PeopleOnStreet		Cactus	
Algoritmo	PSNR	Bitrate	PSNR	Bitrate	PSNR	Bitrate	PSNR	Bitrate
<b>MIRS</b>	37,5701	1706,24	33,4339	3916,88	34,4240	9379,26	35,9173	794871,2
<b>EPZS</b>	37,5642	1704,85	33,3790	3966,24	34,4292	9346,84	36,8187	684160,2
<b>FSS</b>	37,5392	1705,61	33,3491	3976,30	34,3506	9581,12	35,7905	816482,7
<b>DS</b>	37,5473	1707,04	33,3594	3977,08	34,3893	9424,93	26,8304	1244734,8

A partir da Tabela 1 podemos observar que nossas conclusões estavam corretas em trabalhos prévios se tratando de vídeos em qualidade até Full HD. O algoritmo MIRS melhorou a qualidade em 0,3 dB na média em relação ao algoritmo estado-da-arte com a variação de volume de dados em torno de  $\pm 1\%$ , o que podemos considerar como um empate por ser um percentual baixo. Em relação ao FSS e DS o ganho é absoluto em ambos os vídeos nos dois eixos. Isso comprova as perspectivas de ganho para a resolução que o algoritmo foi projetado e tem o mesmo comportamento para baixa resolução.

Observando o comportamento quando aumentamos a resolução para WQXGA temos o EPZS chegando a uma qualidade 0,005 dB melhor para um Bitrate similar. Se aumentarmos para UHD 4K, que é quatro vezes maior que Full HD, a qualidade do EPZS tem uma melhora significativa, chegando a 0,9 dB de vantagem, e uma redução de 16% no Bitrate. Apesar desta perda para o algoritmo de referencia, o MIRS obtém resultados bastante satisfatório se comparado com algoritmos bem conhecidos da literatura como o FSS e DS. Em Bitrate obtém 4% e 36% de redução no volume, e em PSNR 0,13 dB e 9,08 dB.

Para nos aproximar dos resultados ótimos em UHD já esta sendo estudadas novas estratégias. Uma delas é fortificar o algoritmo com um preditor de correlação temporal e espacial que armazene alguns blocos potenciais. Outra estratégia seria construir uma matriz adaptativa, baseada em algoritmos genéticos, que identifique regiões de alta probabilidade de ter bons resultados.

#### 4. CONCLUSÕES

Neste trabalho apresentamos um estudo inicial para o aumento de resolução de Full HD para UHD 4K e seus impactos em qualidade e Bitrate. Os resultados mostram que nossa estratégia atacou bem todos os problemas em Full HD, onde o algoritmo desenvolvido supera em qualidade o algoritmo de referencia EPZS com Bitrate constante. Entretanto, quando aumentamos a resolução para quatro vezes o tamanho de um vídeo Full HD, chegando a resolução UHD 4K, novas características são inseridas na busca, fazendo com que o MIRS seja simplificado para tratar estes vídeos. Ainda assim, comparado com algoritmos clássicos da literatura, o MIRS atinge resultados melhores em qualidade e volume de dados. Além disso, o MIRS tem uma estrutura regular que pode ser processada em paralelo e facilmente sincronizada, beneficiando assim tanto o projeto em Hardware quanto o processamento distribuído em Software.

Como trabalhos futuros iremos propor novas estratégias de busca como a predição por histórico temporal e espacial, e o histórico com matriz adaptativa baseada em algoritmos genéticos. Também será feita a verificação das novas técnicas propostas atuando junto com nossas estratégias consolidadas para verificar uma possível intersecção de resultados. Por fim, esperamos chegar próximo dos resultados ótimos de qualidade para vídeos em UHD 4K e 8K, com algoritmos que facilite o hardware e a exploração paralela.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

IDATE. World Television Market. **IDATE News 541**, 2011.

CISCO. Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. **Cisco Public**, 2013.

KUHN, P., Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation. **Boston: Kluwer Academic Publishers**, 1999.

SULLIVAN, G.; Ohm, J.; et. Al.; " Overview of the High Efficiency Video Coding (HEVC) Standard". **IEEE Transactions on Circuits and Systems for Video Technology**, v. 22, n. 12, 2012.

PORTO, M.; Cristani, C.; et. Al. Iterative Random Search: A New Local Minima Resistant Algorithm for Motion Estimation in High Definition Videos. **Multimedia Tools and Applications**, v. 63, p. 120-127, 2012.

CRISTANI, C.; Dall'Oglio, P.; et. Al. Galaxy Random Search: Algoritmo e Arquitetura para Estimacão de Movimento em Vídeos de Alta Definição. **Revista de Iniciação Científica**, v. 11, 2012.

DALL'OGGIO, P.; Cristani, C.; et. Al. A High Quality Hardware Friendly Motion Estimation Algorithm Focusing in HD Videos. **19th IEEE International Conference on Electronics, Circuits, and Systems**, p. 572-575, 2012.