

PROJETO DE HARDWARE DEDICADO PARA O SOMADOR IMPRECISO LOWER part OR ADDER

MATEUS TERRIBELE LEME; LUCIANO ALMEIDA BRAATZ; LUCIANO
AGOSTINI; MARCELO PORTO

Universidade Federal de Pelotas – Video Technology Research Group (ViTech)
{mtleme, la.braatz, agostini, porto }@inf.ufpel.edu.br

1. INTRODUÇÃO

Certos problemas científicos são resolvidos através da exploração de dois paradigmas de computação: *hard-computing* e *soft-computing*. Existem problemas que exigem altos níveis de precisão numérica para serem solucionados, nestes casos técnicas de *hard-computing* são utilizados por possuírem rigorosos métodos que suportem tal nível de precisão. Essas técnicas são pertinentes nas soluções de alguns problemas de engenharia que demandam uso de modelos e algoritmos determinísticos e precisos (RABAEY, 2010). A precisão oferecida pelos métodos baseados em *hard-computing* detém características de implementação que requerem altos custos computacionais.

Por outro lado, encontram-se problemas resolvidos por *soft-computing*, certa vez que são muito comuns na maioria das aplicações e questão do mundo real, visto que em sua grande parte são naturalmente imprecisos e estocásticos. Desta forma os métodos que visam explorar tolerância, incerteza e verdades parciais, são explorados usando *soft-computing* para alcançar robustez, adaptabilidade e baixo custo (RABAEY, 2010). Incorporado ao paradigma de *soft-computing* existe uma área de atuação chamada computação aproximada que por sua vez possui uma definição equivalente, onde o compromisso é incluir técnicas que expõe intencionalmente o sistema à imprecisão.

A computação aproximada é uma forma de *soft-computing*, que utiliza técnicas que expõe o sistema à imprecisão intencionalmente. Alguns trabalhos desenvolvidos no design de baixa potência através da computação aproximada nos níveis de algoritmo e arquitetura incluem *Algorithmic Noise Tolerance* (SHIM, 2001), *Significance Driven Computation* (MOHAPATRA, 2009) e *Voltage Over-Scaling* (CHAKRAPANI, 2008). Além disso, MAHDIANI (2010) propõe o conceito *Bio-inspired Imprecise Computational Blocks*,

Ilustrado na Figura 1 está o circuito do *Lower part Or Adder (LOA)* incorporado nos blocos imprecisos propostos por MAHDIANI (2010), sua função é computável é $F(A, B) = A + B$, sendo A e B dois números de entrada com largura de bit de tamanho j . A parcela precisa tem j . Ambas as larguras dos números são separadas em duas partes de tamanho i e j , sendo a soma das partes igual à largura do número. Desta forma é possível tratar a soma em duas etapas: precisa e imprecisa. Para computar a parte precisa é realizada a soma utilizando a arquitetura de um somador completo nos bits mais significativos de $j-1$ até i . E a parte imprecisa é computada aplicando a função lógica OU nos bits menos significativos de $i-1$ até 0, assim gerando o restante dos bits e completando a operação de soma. O bit *carry* de entrada do somador completo é gerado por uma função lógica E nos bits mais significativos da parte imprecisa.

Outro cenário de aplicação de *soft-computing* deriva da atual demanda que o processamento das técnicas de compressão de vídeo está exigindo. A codificação de vídeo é um exemplo de aplicação que pode ser melhorada, em termos de eficiência energética, inserindo técnicas de computação aproximada. A inserção de

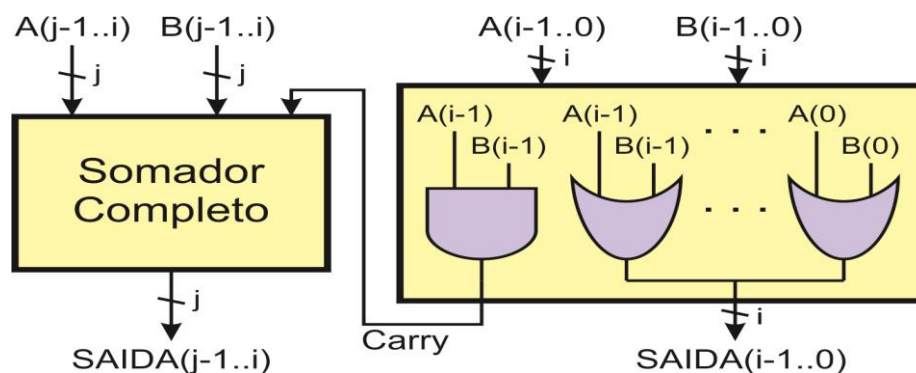


Figura 1- Hardware LOA

computação aproximada nos algoritmos de codificação de vídeo resulta em mudanças imperceptíveis na qualidade do vídeo devido as limitações do sistema visual humano (GAO, 2010). Dessa forma é propício o emprego de computação aproximada em codificadores de vídeo para obter ganhos em velocidade de processamento e redução de energia. Além de que, os vídeos digitais estão em ascensão em uma série de dispositivos, como televisores digitais, smartphones, Blue-ray player, etc. Em virtude desta ascensão e demanda, conveniu-se a utilização de codificadores e decodificadores de vídeo que exploram redundâncias nos dados dos vídeos, assim permitindo uma forma eficiente de armazenar e transmitir os vídeos em tempo real. Atualmente o estado-da-arte em codificação de vídeo é o padrão *HEVC* (JCT-VC, 2013), possui como foco alcançar níveis de compressão cerca de 50% mais altos comparado com seu antecessor *H.264/AVC* (ITU, 2005), e principalmente mantendo a mesma qualidade do vídeo.

O padrão *HEVC* dispõe de etapas com grande complexidade computacional, com isso o desenvolvimento de arquiteturas de *hardware* dedicadas é necessário para o processamento destas etapas em dispositivos com restrições energéticas. O foco deste trabalho está na etapa de codificação, onde são aplicadas dois tipos de transformadas e suas respectivas inversas, a transformada discreta do cosseno (*DCT*) e a transformada discreta do seno (*DST*). Através desta etapa as amostras processadas são transformadas do domínio espacial para o domínio da frequência, que concentra a informação nas baixas frequências, o que reduz a entropia dos dados e facilita a compressão.

A operação de transformação da *DCT* utilizada no *HEVC* é a *DCT-2D* definida em (1), onde N e M são o número de pontos da *DCT*, $F(u, v)$ é a entrada para a posição (u, v) da matriz de entrada e $f(i, j)$ é o coeficiente de saída pertencente a matriz de saída. Uma abordagem utilizada para simplificar a *DCT-2D* consiste na propriedade da separabilidade (GHANBARI, 2003). A separabilidade compõe-se na aplicação de duas *DCT-1D* sucessivamente para obter a *DCT-2D*, dessa maneira diminuindo o número de operações de somas e multiplicações na transformada. Na Tabela 1 Está organizada a quantidade de operações para cada tamanho de bloco de entrada na *DTC-2D*. Em virtude do elevado número de operações nesta etapa de codificação, viabiliza-se a exploração de novos *hardwares* para otimizar o

Tabela 1 - Operações DCT-2D

Tamanho	Somas	Multiplicações
4x4	24	32
8x8	112	128
16x16	480	512
32x32	1984	2048

$$F(u, v) = \sqrt{\frac{2}{N}} \sqrt{\frac{2}{M}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} A_i A_j \cos \left[\frac{\pi u}{2N} (2i + 1) \right] \cos \left[\frac{\pi v}{2M} (2j + 1) \right] f(i, j) \quad (1)$$

desempenho da operação de transformação no HEVC. Uma possível vertente de exploração são os hardwares desenvolvidos sobre arquiteturas baseadas em computação aproximada.

No âmbito de projetos de circuitos integrados, um projetista pode utilizar somadores e multiplicadores imprecisos que são intencionalmente construídos para fornecer uma aproximação do resultado computacional (MAHDIANI, 2010). Esta aproximação pode implicar na redução nos custos de implementação em termos de área, atraso e consumo energético comparado com as arquiteturas precisas. Neste trabalho o objetivo é fazer uma comparação da redução energética e ganhos em velocidade de processamento com diferentes níveis de imprecisão do somador LOA.

2. METODOLOGIA

Para realização deste trabalho a seguinte metodologia foi adotada. Primeiramente, foi realizado a coleta de um conjunto de valores correspondentes às entradas das operações de soma pertencentes ao laço da transformada. Esta coleta de dados foi feita utilizando o software de referência *HEVC Test Model HM*, na versão 16.7. Foi utilizado o vídeo Kimono, de resolução HD 1080p, na configuração Intra e parâmetro de quantização igual a 22.

Posteriormente, foi descrito em VHDL a arquitetura do somador impreciso LOA de 32 bits. Atribuindo diferentes tamanhos da largura de bit na parte imprecisa (0,4,6,8,10,12,14), foi feita a síntese em ASIC para obter uma estimativa de consumo e frequência de operação. Os dados coletados na primeira parte da metodologia são utilizados como entradas para a estimativa de consumo energético, com isso, dados mais confiáveis podem ser obtidos.

3. RESULTADOS E DISCUSSÃO

Os resultados apresentados na Tabela 2 expõe os resultados divididos pela largura da imprecisão, todos os valores foram estimados utilizando cento e vinte mil operações de somas resultantes do processo de aquisição de dados. De maneira mais objetiva é possível identificar comportamentos. O primeiro é a troca inversamente proporcional da precisão (IMP) por frequência (F_MAX), enquanto a precisão diminui a frequência máxima de operação cresce, dessa forma o somador LOA pode funcionar em frequências mais altas. Para obter os ganhos relativos ao consumo energético (G_ENER), foi atribuído a frequência de operação do somador completo às arquiteturas imprecisas e observou um decréscimo na potência dissipada (P_DISSP). A estimativa do erro médio (M_ERRO) foi feita sobre

Tabela 2 - Resultados das sínteses

	IMP0	IMP4	IMP6	IMP8	IMP10	IMP12	IMP14	IMP16
TX_AC(%)	100%	41,3%	25,44%	17,7%	14,1%	13,2%	13,1%	0,37%
M_ERRO	0	11,17	45,05	155,14	565,16	1124,3	1865,3	3690,5
P_DISSP(mW)	0,941	0,922	0,912	0,903	0,889	0,88	0,872	0,863
F_MAX (GHz)	0,986	1,11	1,14	1,21	1,34	1,46	1,55	1,7
G_FREQ(%)	0%	12,7%	15,1%	22,6%	36,1%	48,5%	56,7%	72,5%
G_ENER(%)	0%	2%	3,1%	4%	5,5%	6,5%	7,3%	8,3%

somatório de todos os erros fracionado pelo o número total de operações. O erro médio apresentou outro comportamento associado com a taxa de acerto (TX_AC), por exemplo, com a imprecisão de 4 bits (IMP4) foi de 41,3% a quantidade de acertos, entretanto o erro médio dos resultados imprecisos foram de 11,17.

4. CONCLUSÕES

Neste trabalho foi apresentado um estudo inicial para a redução da complexidade no módulo de transformação do padrão *HEVC* utilizando somadores imprecisos. Como os ganhos em frequência se apresentaram promissores em relação às taxas de erro, é viável explorar a utilização de somadores imprecisos dentro do codificador de vídeo *HEVC*.

Como trabalhos futuros, pretende-se corroborar para vídeos *HD* (1920x1080) e *UHD* (3840x2160) o uso de somadores imprecisos dentro do módulo de transformação. Para realizar este trabalho será necessário o levantamento do impacto da imprecisão na qualidade dos vídeos e desenvolvimento de uma arquitetura imprecisa do módulo. Após essas análises é concebível afirmar o real impacto da imprecisão inserida no módulo de transformação, para determinadas arquiteturas de somadores imprecisos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

J. M. Rabaey, A. Chandraksan, and B. Nikolic. Digital Integrated Circuits, A Design Perspective. **Prentice Hall**, Englewood Cliffs, NJ, p. 591–594, 2003.

B. Shim, S. Sridhara, and N. Shanbhag. Reliable low-power digital signal processing via reduced precision redundancy. **IEEE Trans. VLSI System**, v.12, n. 5, p. 497–510, 2004.

D. Mohapatra, G. Karakonstantis, and K. Roy. Significance driven computation: A voltage-scalable, variation-aware, quality-tuning motion estimator. **Proc. IEEE/ACM International Symposium, on Low Power Electronics and Design**, p. 195–200, 2009.

L. N. Chakrapani, K. K. Muntimadugu, L. Avinash, J. George, and K. V. Highly energy and performance efficient embedded computing through approximately correct arithmetic: a mathematical foundation and preliminary experimental validation. **CASES**, p. 187–196, 2008.

H. R. Mahdiani, A. Ahmadi, S. M. Fakhraie, and C. Lucas. Bio-Inspired Imprecise Computational Blocks for Efficient VLSI Implementation of Soft-Computing Applications. **IEEE Transaction On Circuits And Systems—I: Regular Papers**, v. 57, n. 4, 2010.

GHANBARI, M. Standard Codecs: Image Compression to Advanced Video Coding. United Kingdom. **The Institution of Electrical Engineers**, 2003.

JCT-VC Editors, **Recommendation ITU-T H.265 -High Efficiency Video Coding** (ITU-T Rec.H.265), abril 2013. online. Disponível em: <http://handle.itu.int/11.1002/1000/11885>

International Telecommunication Union (ITU). **ITU-T Recommendation H.264/AVC: advanced video coding for generic audiovisual services**. Janeiro 2012. online. Disponível em: <http://handle.itu.int/11.1002/1000/11466>

X. Gao, W. Lu, D. Tao, X. Li. Image quality assessment and human visual system. **SPIE Video Communications and Image Processing**, v. 7744, p. 77440Z-1-77440Z-10, 2010.