

Funcionalidades do MPEG-5 - Part 2

Low Complexity Enhancement Video Coding (LCEVC)

Por Carlos Eduardo Cosme Ribeiro



Foto [CardMapr.nl](#) de [Unsplash](#)

1. Introdução

Um codificador de vídeo tem como objetivo primário a compressão de um sinal de vídeo para um formato que permita o eficiente armazenamento e/ou transporte do conteúdo comprimido em determinada mídia. Se no passado existiam poucas opções, atualmente existem diversas opções de codificadores de vídeo, cada um com suas peculiaridades, pontos fracos e pontos fortes. O codificador MPEG-2 (ISO/IEC 13818-2, ITU-T H.262), lançado em 1995, foi componente fundamental para viabilizar o nascimento das primeiras gerações de TV Digital no mundo.

Posteriormente, ele foi substituído pelo codificador AVC (*Advanced Video Coding*, ISO/IEC 14496-10, ITU-T H.264), que apresenta maior eficiência de compressão e, mesmo após vinte anos desde o seu lançamento, continua sendo o codificador de vídeo mais utilizado pela indústria de *Broadcast* e pelos atuais serviços de *streaming* [1]. Lançado no ano de 2013, o codificador HEVC (*High Efficiency Video Coding*, ISO/IEC 23008-2, ITU-T H.265), mesmo sendo mais eficiente que o da geração anterior, o AVC, não foi capaz, até o presente momento, de repetir o sucesso do codificador AVC.

O ano de 2020 presenciou o lançamento do substituto do codificador HEVC, isto é, o codificador de vídeo VVC

(*Versatile Video Coding*, ISO/IEC 23090-3, ITU-T H.266), com maior eficiência de compressão e cuja expectativa é repetir o sucesso alcançado pelo AVC. Todos esses codificadores de vídeo foram desenvolvimentos conjuntamente realizados pelos grupos de especialistas de codificação de vídeo da ITU-T (*International Telecommunication Union Telecommunication*) e da ISO/IEC (*International Organization for Standardization/International Electrotechnical Commission*), onde o objetivo primário sempre foi obter um novo codificador de vídeo mais eficiente no quesito compressão de vídeo, que o da geração anterior. No grupo de especialistas da ISO, o MPEG (*Moving Picture Experts Group*), identificou-se a necessidade de se explorar outros paradigmas de forma a abordar questões relacionadas a administração da propriedade intelectual das tecnologias componentes de um codificador de vídeo e, também, a redução da complexidade de codificação/decodificação associado a um baixo consumo energético. O trabalho exploratório abordando o primeiro paradigma resultou no codificador EVC (*Essential Video Coding*, MPEG-5 Part 1 (ISO/IEC 23094-1), lançado em 2020, enquanto a abordagem pelo segundo paradigma resultou no codificador LCEVC MPEG-5 Part 2 (ISO/IEC 23094-2), lançado em 2021 e que é o foco do presente artigo.

2. Motivação do LCEVC

A referência [2] é o *Call for Proposals* emitido pelo MPEG para criação do LCEVC. O objetivo por trás da criação do LCEVC inclui a construção de um codificador de vídeo capaz de agir como um “*enhancer*” para os codificadores de vídeo existentes. Os casos de usos visados incluem a codificação e decodificação em tempo real, alta qualidade de vídeo e retrocompatibilidade com o ecossistema de vídeo, sem a necessidade de atualização e/ou mudança de componentes de hardware. As aplicações alvos iniciais incluem os serviços de *streaming* (ao vivo ou VoD), sistemas de radiodifusão terrestre (enriquecimento de SD para HD e HD para UHD sem necessariamente acarretar um troca total da planta de *set-top-boxes* existentes), *feeds* de vídeos em sistemas de vigilâncias e sistemas de vídeo imersivo.

Para se alcançar tais objetivos, foi elaborada uma arquitetura constituída por uma assim chamada camada base (*base layer*) e uma ou mais camadas de enriquecimento (*enhancement layer*). Na camada base é empregado um codificador de vídeo existente (por exemplo: AVC, HEVC, VVC etc.), porém operando em uma resolução espacial inferior à resolução espacial final desejada. Cabe a uma ou mais camadas de enriquecimento o aprimoramento do produto gerado pela camada base de forma a se obter o sinal de vídeo na resolução espacial final desejada.

O fato da camada base operar com um codificador de vídeo existente, permite ao LCEVC manter retrocompatibilidade com os dispositivos que suportam apenas a tecnologia do codificador de vídeo utilizado na camada base. Observe que, como o codificador de vídeo da camada base opera numa resolução espacial menor, por consequência a sua complexidade é menor, assim como o tempo necessário para codificação do conteúdo do “*base layer*”. Portanto, esses pontos tendem a impactar positivamente na performance global do codificador LCEVC (eficiência de compressão, latência e consumo de energia). Cabe a camada de enriquecimento do LCEVC aprimorar o produto da camada base de forma a se atingir a resolução espacial final desejada e com uma melhor qualidade visual. Na construção desta arquitetura que culminou com a criação do LCEVC, se vislumbrou a camada base sendo implementada via hardware, enquanto as camadas de enriquecimento sendo implementadas via software. Outro importante objetivo desta arquitetura foi torná-la mais eficiente (qualidade e performance) se comparada ao uso individual do codificador da camada base na resolução espacial final desejada. Os trabalhos envolvendo o desenvolvimento do LCEVC ocorreram no MPEG no período de Abril de 2019 até Outubro de 2020 e finalizaram com a publicação da norma do LCEVC em Novembro de 2021.

3. Funcionamento do LCEVC

As figuras 1 e 2 ilustram, respectivamente, os diagramas em blocos funcionais de um codificador LCEVC e de um decodificador LCEVC.

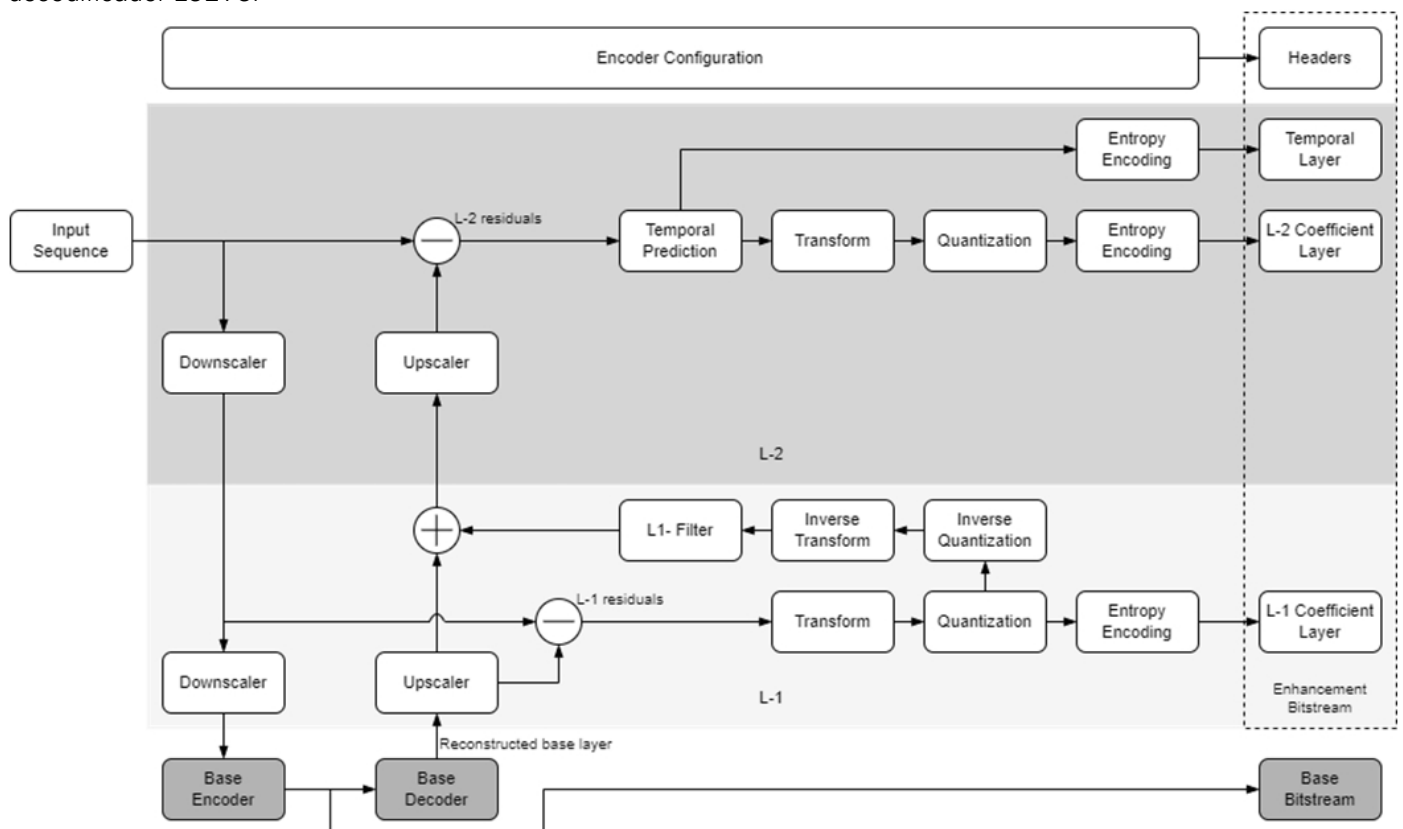


Figura 1: Diagrama em blocos do codificador LCEVC. (Autor)

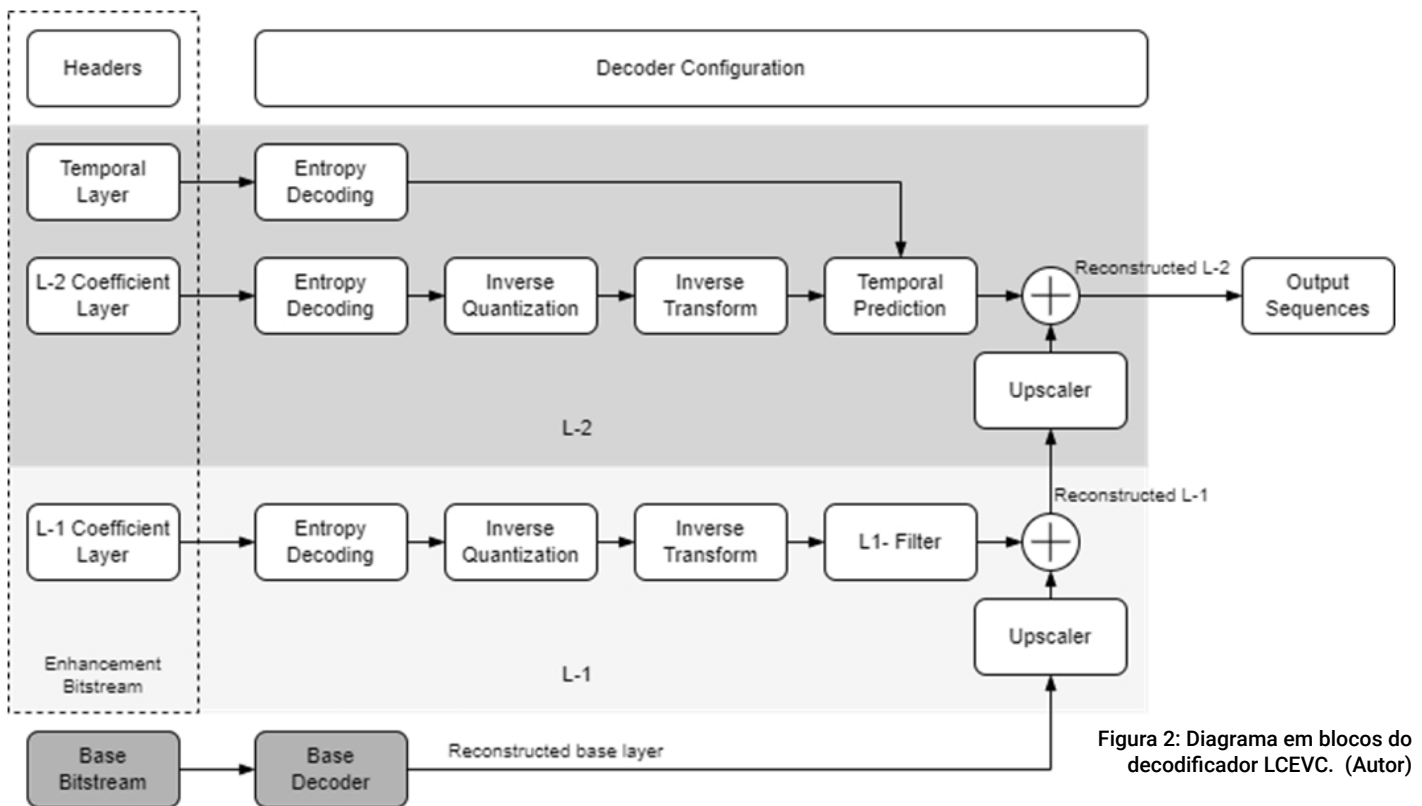


Figura 2: Diagrama em blocos do decodificador LCEVC. (Autor)

O codificador LCEVC é capaz de empregar um máximo de duas camadas de enriquecimento (L-1, *Enhancement layer 1* e L-2, *Enhancement layer 2*) associadas à camada base (*Base layer*). A camada de enriquecimento L-2, além de possuir uma etapa para enriquecimento da resolução temporal, também possui a peculiaridade de sempre estar presente no codificador/decodificador LCEVC. O conteúdo de entrada (*Input sequence*) pode sofrer até dois processos de *downscaling* antes de atingir o *Base Encoder*. Na figura 1, cada *Downscaler* (e correlato *Upscaler* realizando o processo reverso) presentes nas camadas L-1 e L-2, podem através das configurações do LCEVC operarem (1) de forma a manter a largura do sinal de vídeo e reduzir pela metade apenas a altura do sinal de vídeo, (2) reduzir pela metade tanto a altura quanto a largura do sinal de vídeo, ou (3) estarem desabilitados.

Na primeira opção existe a ruptura do quesito retrocompatibilidade dado que a relação de aspecto do produto da camada base será diferente da relação de aspecto do produto final reconstruído, enquanto que nas demais opções o produto da camada base e o produto da camada de enriquecimento possuem a mesma relação de aspecto. A diferença operacional entre as opções 2 e 3 reside no fato de que na opção 2 o enriquecimento agrega um incremento da resolução espacial, por exemplo, poderíamos ter uma camada base em HD, um primeiro enriquecimento aumentando a resolução espacial para 4K e um segundo enriquecimento aumentando a resolução espacial para 8K. Na opção 3, os sucessivos enriquecimentos oferecidos apenas incrementarão a qualidade visual do conteúdo base dado que a resolução

espacial do conteúdo permanece imutável ao longo do processamento nas camadas base e de enriquecimento (*Downscalers* e *Upscalers* desligados).

O *Base Encoder* basicamente segue o processo de codificação de acordo com o codificador base escolhido (AVC, HEVC, VVC etc.). A grande diferença no codificador LCEVC ocorre nos processos de codificação dentro das camadas de enriquecimento L-1 e L-2. Observe pela figura 1 que o processo de codificação LCEVC nas camadas de enriquecimento sempre ocorre sobre o conteúdo residual obtido da diferença entre o produto de entrada e o produto reconstruído em dada camada de enriquecimento. A esse conteúdo residual é aplicado um processo de transformação, cujo tamanho de bloco pode ser de 2x2 ou 4x4 pixels, contribuindo assim para a manutenção dos detalhes (informação de alta frequência) em um conteúdo visual. Na sequência, os coeficientes da transformada são quantizados de forma linear e seguem para o processo de codificação de entropia, composto por dois componentes: *Run Length Encoder* e *Prefix Encoder*. Após esta etapa, a informação codificada estará disponível para ser encapsulada nas unidades NAL (*Network Abstraction Layer*) e efetivamente inseridas nos container da tecnologia de transporte utilizada (MPEG-TS ou ISO/BMFF). Exclusivamente, na camada de enriquecimento L-2 existe um bloco de previsão temporal que utiliza a técnica de previsão temporal chamada *zero-motion vector*, onde a informação de vídeo residual do frame atual é processada junto com a informação de vídeo residual do frame anterior de forma a se determinar o vetor de movimento.

As informações das camadas de enriquecimento (*Enhancement bitstream*) podem ser transportadas em uma das seguintes formas:

- Encapsuladas no mesmo container da camada base (*Base bitstream*) por meio do emprego de mensagens SEI (*Supplemental Enhancement Information*). Por exemplo, no caso do MPEG Transport Stream estariam contidas dentro do PID (*Packet Identifier*) de vídeo responsável pelo transporte das informações de Base bitstream referentes ao codificador base.
- Encapsuladas em um container diferente do container utilizado para transporte das informações Base bitstream mas utilizando a mesma tecnologia de transporte. Por exemplo, no caso do ISO BMFF (*ISO Base Media File Format*) as informações de Base bitstream são encapsuladas em determinados boxes

MDAT (*Media Data box*), enquanto as informações Enhancement bitstream são encapsuladas em diferentes boxes MDAT. Na sequência, os dois conjuntos de boxes MDAT podem ser enviados juntos pelo mesmo canal de distribuição (por exemplo, *over-the-air* ou *over-the-top*) ou serem distribuídos por diferentes canais de distribuição (por exemplo, *over-the-air* e *over-the-top*).

O decodificador LCEVC ilustrado na figura 2 funciona de forma inversa ao codificador LCEVC, reconstruindo o produto da camada base e o posteriormente enriquecendo-o com uma ou mais camadas de enriquecimento até se obter o conteúdo em sua forma final (*Output sequences*). Para maiores informações e detalhes específicos do funcionamento do LCEVC, recomenda-se a leitura da base de informações disponível no website da empresa V-Nova [7].

4. Aplicações com LCEVC

O codificador de vídeo LCEVC é uma tecnologia desenvolvida pelo MPEG mas cuja patente é detida pela empresa V-Nova sediada em Londres, Inglaterra. Conforme mencionado anteriormente, o LCEVC permite o enriquecimento do sinal de vídeo e ainda a manutenção da retrocompatibilidade com dispositivos que apenas suportam a tecnologia de codificação de vídeo empregada em sua camada base. Um outro caso de uso recentemente explorado com o codificador LCEVC, é permitir uma camada base cuja a profundidade de bits, a taxa de quadros, a curva de transferência e o espaço de cores, sejam diferentes aos obtidos após o enriquecimento com as informações das camadas de enriquecimento. As referências [5] e [6] reportam uma experimentação referente a este caso de uso, realizada durante a Copa do Mundo de 2022. Nesta experimentação, o sinal original do evento esportivo no formato 1080p, 59.94 fps, espaço de cores BT.2020, curva de transferência HDR10 alimentou um encoder experimental LCEVC+H.264 (LCEVC com H.264 como codificador da camada base). O referido encoder produziu uma informação de Base layer no formato 1080i, 29,97 fps, espaço de cores BT.709, curva de transferência BT.709 e informações da camada de enriquecimento que aprimoraram a qualidade da camada base de volta ao formato original na entrada do encoder.

Outra aplicação interessante para o codificador LCEVC é na geração do *bitstream ladder* dos serviços de *streaming*. Por exemplo, um determinado serviço de *streaming* que contenha uma *ladder* H.265 SDR (*Standard Dynamic Range*) visando atender dispositivos legados

e outra *ladder* H.265 HDR (*High Dynamic Range*) para atender dispositivos recentes; poderia empregar apenas uma única *ladder* com o codificador LCEVC associado ao codificador H.265. Nesse caso, a camada base empregaria o formato H.265 SDR (para compatibilidade com dispositivos legados) e as camadas de enriquecimento melhorariam o produto da camada base visando-se uma qualidade HDR final semelhante ou até mesmo superior à obtida se comparada ao emprego do codificador H.265 sozinho. A magnitude da taxa de bits para a camada de enriquecimento é um parâmetro que a depender da configuração empregada, oscila em 10 a 50% da taxa de bits empregada na camada base.



Foto de [mahmoud azmy](#) em [Unsplash](#)

5. Conclusões

Conforme apresentado no artigo, o codificador LCEVC é um codificador que oferece diversas possibilidades no seu emprego e também é uma das tecnologias de vídeo escolhidas para a futura TV 3.0 no Brasil. No projeto

TV 3.0, o codificador LCEVC terá a responsabilidade de incrementar a qualidade do conteúdo da camada base (que será codificado com a tecnologia VVC).

Referências

- [1] The 6th Annual Bitmovin Video Developer Report, Shaping the future of video 2022/2023. Disponível em <https://bitmovin.com/video-developer-report/#pdf>
- [2] MPEG Requirements for LCEVC, MPEG 142 meeting. Disponível em <https://mpeg.chiariglione.org/standards/exploration/low-complexity-video-coding-enhancements/requirements-low-complexity-video>
- [3] Whitepaper on Low Complexity Enhancement Video Coding (LCEVC), MPEG 137 meeting. Disponível em <https://www.lcevc.org/wp-content/uploads/Whitepaper-on-Low-Complexity-Enhancement-Video-Coding-LCEVC-MPEG-meeting-137-January-2022.pdf>
- [4] MPEG Verification Test Report on the Compression Performance of LCEVC, MPEG 134 meeting. Disponível em <https://www.lcevc.org/wp-content/uploads/MPEG-Verification-Test-Report-on-the-Compression-Performance-of-LCEVC-Meeting-MPEG-134-May-2021.pdf>
- [5] ITU-R SG06.AR Contribution 11, "TV 2.5 and TV 3.0 Trials on the FIFA World Cup 2022", 2023-02-14.
- [6] ISO/IEC JTC 1/SC 29/WG4, m62044, "Report on usage of LCEVC in Brazilian broadcasting for demonstration during Football World Cup", January 2023.
- [7] V-Nova website, LCEVC Resources section. Disponível em <https://www.lcevc.org/lcevc-resources/>



Carlos Eduardo Cosme Ribeiro é graduado em Engenharia Elétrica pela Universidade Estácio de Sá (2007), pós-graduado em Engenharia de Controle e Automação pela Universidade Gama Filho (2011) e especialista em Redes de Computadores pela PUC Rio (2013). Trabalha na Globo desde 1998, onde atuou nas áreas de operação e suporte. Atualmente é especialista em inovação no grupo de Telecom do Hub de Infraestrutura e Segurança. É membro do Fórum do Sistema Brasileiro de TV Digital Terrestre (SBTVD), onde coordena os grupos de trabalho de Codificação de Áudio, Codificação de Vídeo e Legendas no Projeto TV 3.0.

Contato: cecosme@g.globo



NOVO APP

Baixe o SET Connect e fique conectado com a SET.

Notícias, Eventos, Oportunidades e Conteúdos Exclusivos para Associados

