

Padrão de compactação de vídeo H.264.
Novas possibilidades em vigilância por vídeo.

Índice

1. Introdução	3
2. Desenvolvimento do H.264	3
3. Como funciona a compactação de vídeo	4
4. Perfis e níveis do H.264	5
5. Compreensão dos quadros	5
6. Métodos básicos de redução de dados	6
7. Eficiência do H.264	7
8. Conclusão	9

1. Introdução

O mais recente padrão de compactação de vídeo, o H.264 (também conhecido como MPEG-4 parte 10/AVC para codificação de vídeo avançada), deve se tornar o padrão de vídeo escolhido nos próximos anos.

O H.264 é um padrão aberto e licenciado que é compatível com as técnicas mais eficientes de compactação de vídeo disponíveis atualmente. Sem comprometer a qualidade de imagem, um codificador H.264 pode reduzir o tamanho de um arquivo de vídeo digital em mais de 80% em comparação com o formato Motion JPEG e em até 50% mais do que o tradicional padrão MPEG-4 parte 2. Isso significa que muito menos largura de banda de rede e espaço de armazenamento são necessários para um arquivo de vídeo. Ou visto de outra forma, pode ser atingida uma qualidade de vídeo bem mais superior para determinada taxa de bits.

Definido com conjunto pelas organizações de padronização dos setores de telecomunicações e TI, o H.264 deve ser mais amplamente adotado do que os padrões anteriores.

O H.264 já foi introduzido em novos dispositivos eletrônicos, como celulares e reprodutores de vídeo digital, e tem ganhado rápida aceitação pelos usuários finais. Os provedores de serviços, como empresas de telecomunicações e armazenamento de vídeo on-line, também estão começando a adotar o H.264.

No setor de vigilância por vídeo, o H.264 provavelmente terá adesão mais rápida em aplicações nas quais haja demandas para altas taxas por quadro e alta resolução, como na vigilância de estradas, aeroportos e cassinos, onde o uso de 30/25 (NTSC/PAL) quadros por segundo é a regra. É nesse ponto que a diminuição da necessidade de armazenamento e largura de banda reduzida fornecerá as maiores reduções de gastos.

O H.264 também deve acelerar a adoção de câmeras de megapixel, pois a tecnologia de compactação extremamente eficiente pode reduzir os tamanhos de arquivos grandes e taxas de bits geradas sem comprometer a qualidade da imagem. No entanto, há concessões. Ao mesmo tempo que o H.264 possibilita economia nos custos de armazenamento e largura de banda de rede, ele exigirá câmeras de rede e estações de monitoramento de desempenho superior.

2. Desenvolvimento do H.264

O H.264 é o resultado de um projeto conjunto entre o Video Coding Experts Group (VCEG) da International Telecommunications Union (ITU) e o ISO/IEC Moving Picture Experts Group (MPEG). O ITU-T é o setor que coordena os padrões de telecomunicações em nome da International Telecommunication Union. ISO significa International Organization for Standardization (Organização Internacional para Padronização) e IEC é uma organização que supervisiona padrões eletrônicos e elétricos, bem como tecnologias relacionadas. O H.264 é o nome usado pela ITU-T, enquanto a ISO/IEC o denominou MPEG-4 parte 10/AVC, pois ele é apresentado como uma nova parte em sua plataforma MPEG-4. A plataforma MPEG-4 inclui, por exemplo, o MPEG-4 parte 2, que é um padrão que tem sido utilizado pelas câmeras de rede e codificadores de vídeo baseados em IP.

Desenvolvido para solucionar várias deficiências nos padrões de compactação de vídeo anteriores, o H.264 atende suas metas de suporte a:

- > Implementações que fornecem uma redução média de taxa de bits de 50%, com base em uma qualidade vídeo fixa em comparação a qualquer outro padrão de vídeo
- > Robustez de erro, de modo que os erros de transmissão em várias redes sejam tolerados
- > Recursos de baixa latência e melhor qualidade para maior latência
- > Especificação clara de sintaxe que simplifica implementações
- > Decodificação de correspondência exata, que define exatamente como os cálculos numéricos são feitos por um codificador e decodificador, a fim de evitar erros decorrentes de acúmulo

O H.264 também tem a flexibilidade de oferecer suporte a uma ampla variedade de aplicações com requisitos muito diferentes de taxa de bits. Por exemplo, em aplicações de vídeo de entretenimento – que incluem transmissão, satélite, cabo e DVD – o H.264 poderá fornecer um desempenho entre 1 a 10 Mbit/s com alta latência, enquanto para serviços de telecomunicações, o H.264 pode fornecer taxas de bits abaixo de 1 Mbit/s com baixa latência.

3. Como funciona a compactação de vídeo

A compactação de vídeo diz respeito à redução e à remoção de dados redundantes de vídeo, de modo que um arquivo de vídeo digital possa ser efetivamente enviado e armazenado.

O processo envolve aplicar um algoritmo ao vídeo de origem para criar um arquivo compactado já pronto para transmissão ou armazenamento. Para reproduzir o arquivo compactado, um algoritmo inverso será aplicado para produzir um vídeo que mostre praticamente o mesmo conteúdo que o vídeo original. O tempo que demora para compactar, enviar, descompactar e exibir um arquivo é denominado latência. Quanto mais avançado o algoritmo de compactação, maior a latência, devido a mesma capacidade de processamento.

Um par de algoritmos que funcionam juntos é denominado como codec de vídeo (codificador/decodificador). Os codecs de vídeo que implementam diferentes padrões geralmente não são compatíveis entre si; ou seja, o conteúdo de vídeo que é compactado usando um padrão não pode ser descompactado com outro padrão. Por exemplo, um decodificador MPEG-4 parte 2 não funcionará com um codificador H.264. Isso se deve simplesmente ao fato de um algoritmo não poder decodificar corretamente a saída de outro algoritmo, mas é possível implementar muitos algoritmos diferentes no mesmo software ou hardware, o que permitiria a compactação de vários formatos.

Diferentes padrões de compactação de vídeo utilizam diferentes métodos para reduzir dados e, portanto, resultam em diferentes taxas de bits, qualidade e latência. Os resultados de codificadores que usam o mesmo padrão de compactação também podem variar, devido ao fato do designer de um codificador poder escolher implementar diferentes conjuntos de ferramentas definidos por um padrão. Desde que a saída de um codificador esteja em conformidade com um formato e decodificador do padrão, é possível fazer diferentes implementações. Isso é vantajoso, pois diferentes implementações têm diferentes metas e orçamento. Os codificadores de software profissionais que não sejam em tempo real para controle de mídia óptica devem ter a opção de poder fornecer melhor vídeo codificado do que um codificador de hardware em tempo real para conferência de vídeo integrada em um dispositivo portátil. Portanto, determinado padrão não garante determinada qualidade ou taxa de bits. Além disso, o desempenho de um padrão não pode ser comparado corretamente a outros padrões ou até mesmo a outras implementações do mesmo padrão, sem primeiro definir como será implementado.

Um decodificador, ao contrário de um codificador, deve implementar todas as partes necessárias de um padrão a fim de decodificar um fluxo de bits em conformidade. Isso se deve ao fato de um padrão especificar exatamente como um algoritmo de descompactação deve restaurar cada bit de um vídeo compactado.

O gráfico a seguir fornece uma comparação da taxa de bits, em virtude do mesmo nível de qualidade de imagem, entre os seguintes padrões de vídeo: Motion JPEG, MPEG-4 parte 2 (sem compensação de movimento), MPEG-4 parte 2 (com compensação de movimento) e H.264 (perfil da linha de base).

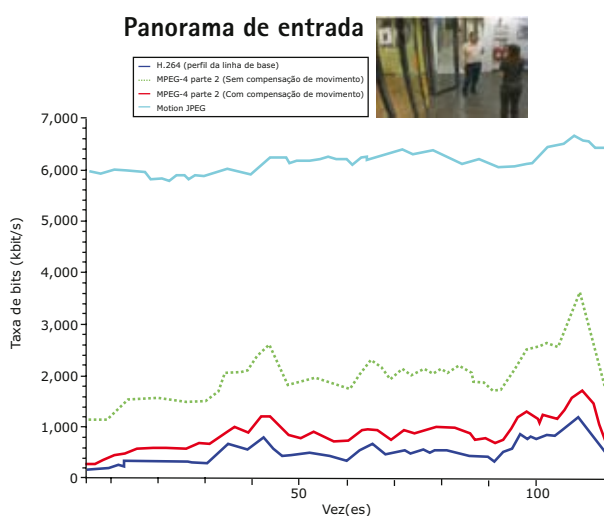


Figura 1. Um codificador H.264 gerou até 50% menos bits por segundo para uma seqüência de vídeo de amostra do que um codificador MPEG-4 com compensação de movimento. O decodificador H.264 foi, pelo menos, três vezes mais eficiente do que um codificador MPEG-4 sem compensação de movimento e, pelo menos, seis vezes mais eficiente do que o Motion JPEG.

4. Perfis e níveis do H.264

O grupo conjunto participou da definição do H.264 com foco na criação de uma solução simples e clara, limitando as opções e os recursos a um mínimo. Um importante aspecto do padrão, como com outros padrões de vídeo, é fornecer os recursos em perfis (conjuntos de recursos algorítmicos) e níveis (classes de desempenho) que idealmente são compatíveis com produções populares e formatos comuns.

O H.264 tem vários perfis, cada um direcionado a uma classe específica de aplicações. Cada perfil define qual conjunto de recursos o codificador pode usar e limita a complexidade de implementação do decodificador.

As câmeras de rede e os codificadores de vídeo provavelmente utilizarão um perfil denominado perfil da linha de base, que se destina principalmente a aplicações com recursos limitados de computação. O perfil da linha de base é mais adequado de acordo com o desempenho disponível em um codificador em tempo real, que é incorporado em um produto de vídeo de rede. O perfil também permite baixa latência, que é um requisito importante de vídeo de vigilância e também especialmente importante para habilitar o controle em tempo real de PTZ (pan/tilt/zoom ou panorama/inclinação/zoom) em câmeras de rede PTZ.

O H.264 tem 11 níveis ou grau de recursos para limitar os requisitos de desempenho, largura de banda e memória. Cada nível define a taxa de bits e a taxa de codificação em macroblocos por segundo para resoluções variando de QCIF a HDTV e outras. Quanto maior a resolução, maior o nível necessário.

5. Compreensão dos quadros

Dependendo do perfil do H.264, diferentes tipos de quadros, como quadros I, quadros P e quadros B, podem ser usados por um codificador.

Um quadro I, ou quadro intra, é um quadro de autocontrole que pode ser decodificado independentemente sem qualquer referência a outras imagens. A primeira imagem em uma seqüência de vídeo é sempre um quadro I. Os quadros I são necessários como pontos de partida para novos visualizadores ou pontos de resincronização, se o fluxo de bits transmitidos for danificado. Os quadros I podem ser usados para implementar avanço rápido, retrocesso e outras funções de acesso aleatório. Um codificador inserirá automaticamente quadros I em intervalos regulares ou mediante solicitação, se novos clientes forem participar da visualização de um fluxo. A desvantagem dos quadros I é que eles consomem muito mais bits, mas por outro lado, não geram muitos artefatos.

Um quadro P, que significa quadro inter de previsão, faz referências a partes de quadros I e/ou P anteriores para codificar o quadro. Os quadros P geralmente exigem menos bits do que quadros I, mas uma desvantagem é que eles são muito sensíveis a erros de transmissão, devido à dependência complexa de quadros de referência P e I anteriores.

Um quadro B, ou quadro inter bipreditivo, é um quadro que faz referências a um quadro de referência anterior e um quadro futuro.

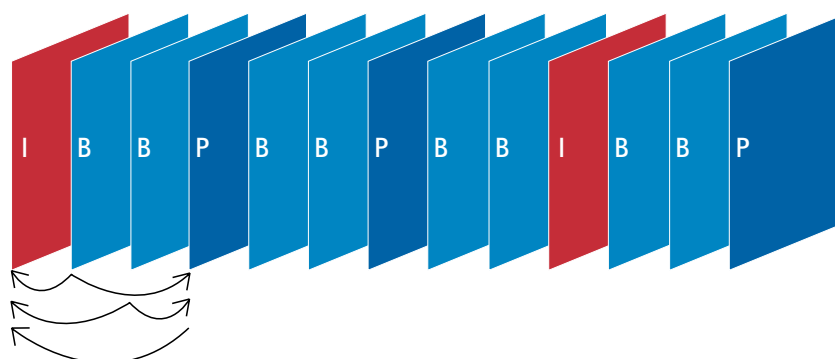


Figura 2. Uma seqüência típica com quadros I, B e P. Um quadro P pode fazer referência somente a quadros I ou P anteriores, enquanto um quadro B pode fazer referência a quadros I ou P anteriores e posteriores.

Quando um decodificador de vídeo restaura um vídeo ao decodificar quadro a quadro de fluxo de bits, a decodificação deve sempre iniciar com um quadro I. Os quadros P e B, se usados, devem ser decodificados juntos com os quadros de referência.

No perfil da linha de base do H.264, somente os quadros I e P são usados. Esse perfil é ideal para câmeras de rede e codificadores de vídeo, pois a baixa latência é atingida já que os quadros B não são usados.

6. Métodos básicos de redução de dados

Uma série de métodos pode ser usada para reduzir os dados de vídeo, em um quadro de imagem e entre uma série de quadros.

Em um quadro de imagem, os dados podem ser reduzidos simplesmente pela remoção das informações desnecessárias, que terão um impacto na resolução da imagem.

Em uma série de quadros, os dados de vídeo podem ser reduzidos por esses métodos como codificação de diferença, que é usado pela maioria dos padrões de compactação de vídeo, incluindo o H.264. Na codificação de diferença, um quadro é comparado com um quadro de referência (p. ex., quadro I ou P anterior) e somente os pixels que tiverem sido alterados em relação ao quadro de referência são codificados. Dessa forma, o número de valores de pixel que são codificados e enviados é reduzido.



Figura 3. Com o formato Motion JPEG, as três imagens na seqüência acima são codificadas e enviadas como imagens únicas separadas (quadros I) sem dependências entre si.

— Transmitido - - - Não transmitido



Figura 4. Com a codificação de diferença (usada na maioria dos padrões de compactação de vídeo, incluindo H.264), somente a primeira imagem (quadro I) é codificada em sua totalidade. Nas duas imagens a seguir (quadros P), são feitas referências à primeira imagem dos elementos estáticos, p. ex., a casa e somente as partes em movimento, ou seja, o homem correndo, são codificadas por meio de vetores de movimento, reduzindo assim a quantidade de informações enviadas e armazenadas.

A quantidade de codificação pode ser ainda mais reduzida se a detecção e a codificação de diferenças forem baseadas em blocos de pixels (macroblocos) em vez de pixels individuais; portanto, as áreas maiores são comparadas e somente os blocos significativamente diferentes são codificados. A sobrecarga associada à indicação da localização de áreas a serem alteradas também é reduzida.

No entanto, a codificação de diferenças não reduziria significativamente os dados se houvesse muito movimento em um vídeo. Aqui, podem ser usadas técnicas como compensação de movimento com base

em blocos. A compensação de movimento baseada em blocos considera quanto do que compõe um novo quadro em uma seqüência de vídeo pode ser localizado em um quadro anterior, mas talvez em um local diferente. Essa técnica divide um quadro em uma série de macroblocos. Bloco por bloco, um novo quadro – por exemplo, um quadro P – pode ser composto ou “previsto” ao verificar um bloco de correspondência em um quadro de referência. Se uma correspondência for localizada, o codificador simplesmente codificará a posição na qual o bloco de correspondência deve estar localizado no quadro de referência. A codificação do vetor de movimento, como é chamada, utiliza menos bits do que se o conteúdo real de um bloco fosse codificado.

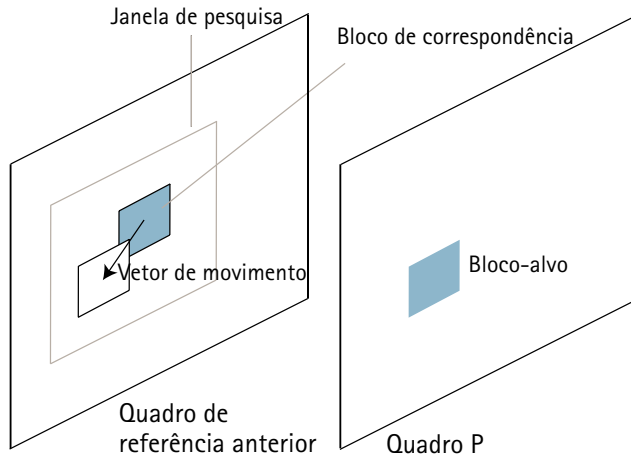


Figura 5. Ilustração da compensação de movimento baseada em bloco.

7. Eficiência do H.264

O H.264 eleva a tecnologia de compactação de vídeo para um novo patamar.

Com o H.264, um novo e avançado esquema de previsão intra é introduzido para codificação de quadros I. Esse esquema pode reduzir significativamente o tamanho de bits de um quadro I e manter uma alta qualidade ao permitir a previsão sucessiva de blocos menores de pixels em cada macrobloco de um quadro. Isso é feito ao tentar localizar pixels de correspondência entre os pixels codificados anteriormente que limitam um novo bloco de pixels 4x4 para ser intracodificado. Ao reutilizar os valores de pixel que já foram codificados, o tamanho de bits pode ser drasticamente reduzido. A nova previsão intra é uma parte fundamental da tecnologia do H.264 que é comprovadamente muito eficiente. Para fins de comparação, se somente os quadros I fossem usados em um fluxo de H.264, ele teria um tamanho de arquivo muito menor do que um fluxo de Motion JPEG, que usa somente quadros I.

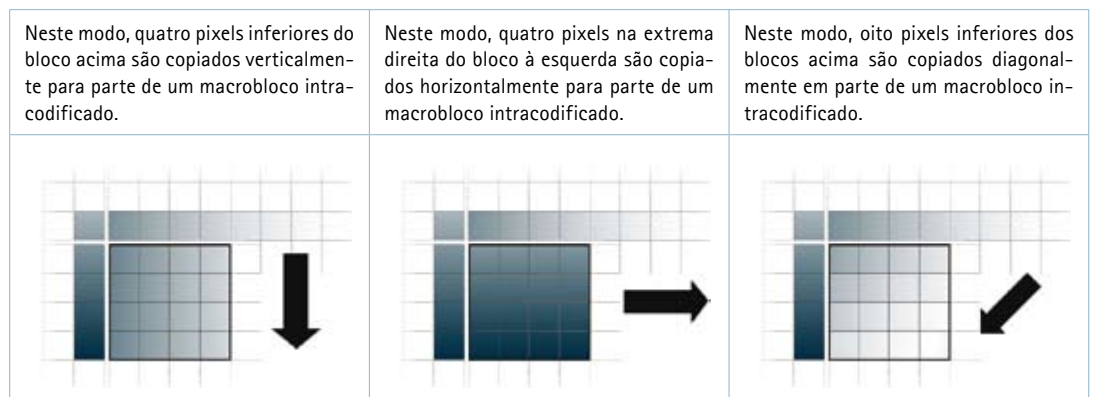


Figura 6. As ilustrações de alguns dos modos com previsão intra podem se considerados na codificação de pixels 4x4 em um dos 16 blocos que compõem um macrobloco. Cada um dos 16 blocos em um macrobloco pode ser codificado usando diferentes modos.



Imagem de fonte original



Imagem de previsão intra

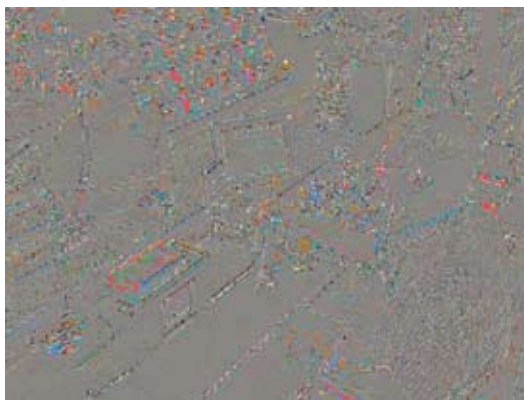


Imagem residual



Imagem de saída

Figura 7. As imagens acima ilustram a eficiência do esquema de previsão intra do H.264, onde a imagem intra prevista é enviada "gratuitamente". Somente os modos de previsão intra e conteúdo residual precisam ser codificados para gerar a imagem de saída.

A compensação de movimento com base em blocos – utilizada na codificação de quadros P e B – também foi aprimorada no H.264. Um codificador H.264 pode optar por pesquisar blocos de correspondência – abaixo para precisão de subpixel – em algumas ou muitas áreas de um ou vários quadros de referência. O tamanho e o formato do bloco também podem ser ajustados para aprimorar uma correspondência. Em áreas em que não há blocos de correspondência em um quadro de referência, são utilizados macroblocos intracodificados. O alto grau de flexibilidade em compensação de movimento com base em bloco do H.264 é válido em cenas de vigilância com muitos elementos, em que a qualidade pode ser mantida para aplicações exigentes. A compensação do movimento é o aspecto mais exigente de um codificador de vídeo e as diferentes formas e graus com os quais ele pode ser implementado por um codificador H.264 podem ter um impacto na eficiência de compactação de vídeo.

Com o H.264, os artefatos geralmente em bloco – vistos em vídeo extremamente compactado usando padrões Motion JPEG e MPEG que não sejam H.264 – podem ser reduzidos usando um filtro de eliminação de blocos em loop. Esse filtro suaviza as bordas de blocos usando uma potência adaptável para fornecer um vídeo descompactado quase que perfeito.



Figura 8. *Os artefatos em blocos em imagens extremamente compactadas à esquerda são reduzidos quando um filtro para eliminar os blocos é aplicado, como visto na imagem à direita.*

8. Conclusão

O H.264 apresenta um grande avanço na tecnologia de compactação de vídeo. Ele oferece técnicas que permitem melhores eficiências de compactação devido aos recursos de previsão mais precisos, bem como resiliência aprimorada a erros. Ele fornece novas possibilidades para criar melhores codificadores de vídeo que permitam fluxos de vídeo de qualidade superior, taxas de quadro mais altas e resoluções superiores a taxas de bits mantidas (em comparação aos padrões anteriores) ou a mesma qualidade de vídeo a taxas de bits inferiores.

O H.264 representa a primeira vez que a ITU, ISO e a IEC se reuniram em um padrão comum e internacional para compactação de vídeo. Devido à sua flexibilidade, o H.264 tem sido aplicado em diversas áreas, como DVD de alta definição (p. ex., Blu-ray), transmissão de vídeo digital, incluindo TV de alta definição, armazenamento de vídeo on-line (p. ex., YouTube), telefonia móvel de terceira geração, em software como QuickTime, Flash e sistema operacional MacOS X da Apple Computer e em consoles de jogos de vídeo domésticos, como o PlayStation 3. Com suporte de várias indústrias e aplicações para necessidades profissionais e de clientes, o H.264 deve substituir outros padrões e métodos de compactação utilizados atualmente.

À medida que formato H.264 se tornar mais amplamente disponível em câmeras de rede, codificadores de vídeo e software de gerenciamento de vídeo, os designers e integradores de sistemas precisarão se certificar de que os produtos e os fornecedores que escolherem ofereçam suporte a esse novo padrão aberto. E, até o momento, os produtos de vídeo de rede compatíveis com H.264 e Motion JPEG são ideais para máxima flexibilidade e possibilidades de integração.

Sobre a Axis Communications

A Axis é uma empresa de TI que oferece soluções para instalações profissionais. A empresa é líder do segmento no mercado global e impulsiona a contínua migração do sistema analógico de vigilância por vídeo para o digital. Seus produtos e soluções, baseados em plataformas de tecnologias novas e abertas, são focados em vigilância e monitoramento remoto.

Com sede na Suécia, a Axis possui escritórios em 18 países e parceiros em mais de 70. Fundada em 1984, está incluída no OMX Nordic Exchange, Large Cap and Information Technology. Para obter mais informações, visite o site www.axis.com