

1

Introdução

1.1

Mudança no paradigma da codificação de vídeo

Com o desenvolvimento de novos dispositivos eletrônicos que embarcam modernas tecnologias (câmeras IP sem-fio, celulares e *smartphones* com acesso multi-redes, *tablets* e laptops com novas tecnologias *wi-fi*, redes de sensores e outros) e o aumento da infra-estrutura de comunicações dando suporte a essas inovações, a informação e os dados digitais passaram a ter a ubiquidade como uma de suas principais características, já que estes podem ser acessados e transmitidos de qualquer lugar. Isso fez com que as arquiteturas tradicionais de transmissão/recepção de dados procurassem novos padrões de disposição e novas arquiteturas, de forma que seu funcionamento se adaptasse a esse cenário, aproveitando a capacidade tecnológica desses dispositivos para a criação de novas soluções, serviços e negócios.

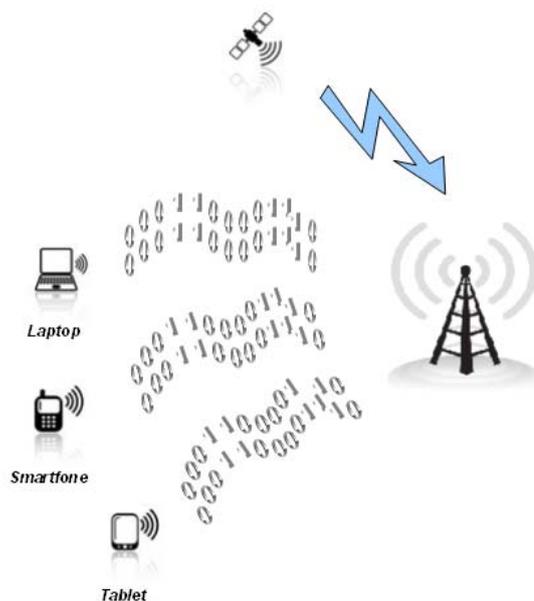


Figura 1.1 – Dispositivos eletrônicos e ubiquidade: acesso a informações através de laptop, *smartphones*, *tablets* e outros dispositivos eletrônicos modernos

Em relação à transmissão de vídeo, devido às características inerentes a esse tipo de informação, a maioria das arquiteturas tradicionais é assimétrica, com a codificação de vídeo realizada por uma máquina de maior complexidade e, após a transmissão e recepção, a decodificação é realizada num dispositivo cliente com operações de menor complexidade. Esse cenário se mostra adequado, devido à necessidade de se ter um aparelho de baixo custo servindo de atrativo ao usuário final, o que deve ser compensado com o volume de dispositivos clientes e com o custo a ser cobrado pelo serviço ou propagandas inseridas no conteúdo. Essa arquitetura também se mostra razoável, na outra ponta, a partir do momento em que se tem um cenário no qual a produção do conteúdo, codificação e transmissão são feitos por empresas com capacidade de comprar equipamentos com grande poder de processamento e garantir ao cliente de seu serviço um baixo custo de acesso a esse conteúdo, através de um aparelho com um decodificador barato e com baixos requisitos computacionais. Isso pode ser exemplificado pela relação que se tem num sistema de transmissão de TV digital terrestre aberta (*broadcasting*), considerando as características técnicas das empresas de televisão e dos clientes telespectadores, como está ilustrado na figura 1.2.

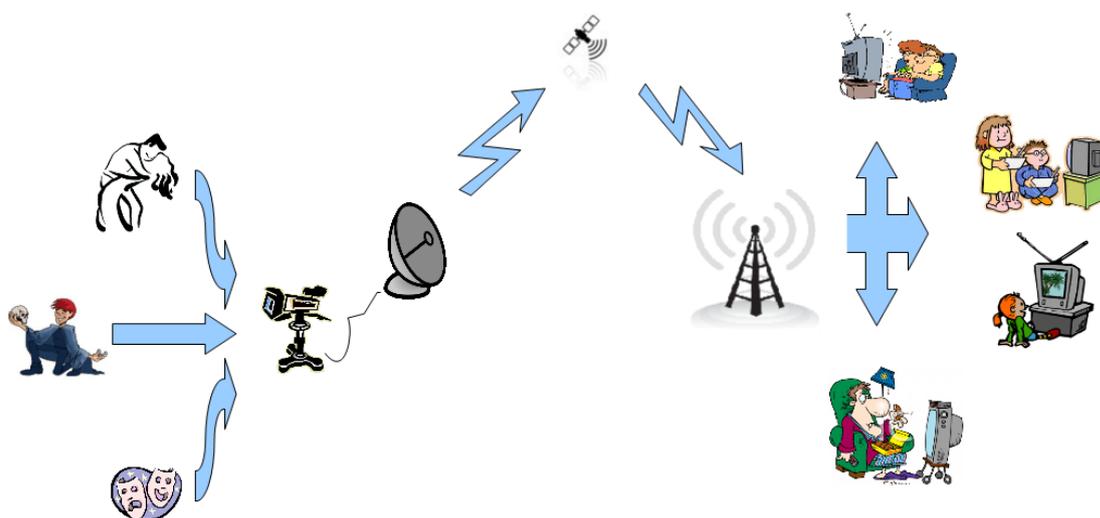


Figura 1.2 – Cadeia de consumo do conteúdo de tv

Diferentemente do que ocorre no cenário anteriormente descrito, novas tecnologias focadas em mobilidade e transmissão, principalmente através de meios não confinados (sem-fio), fizeram com que houvesse necessidade de uma

mudança de paradigma em relação à complexidade e custo técnico dos dispositivos utilizados. O hardware embarcado no codificador tornou-se mais simples, já que passamos a ter aparelhos com requisitos diferentes do cenário anterior considerando-se a produção do conteúdo (onde temos o codificador), como mobilidade, peso, capacidade de processamento, autonomia da bateria entre outras características.

Em contrapartida, a complexidade para quem faz a decodificação foi aumentada, exigindo um equipamento mais caro, com maior capacidade de processamento e, conseqüentemente, habilitado a lidar com uma maior complexidade algorítmica, conforme exemplificado na figura 1.3.

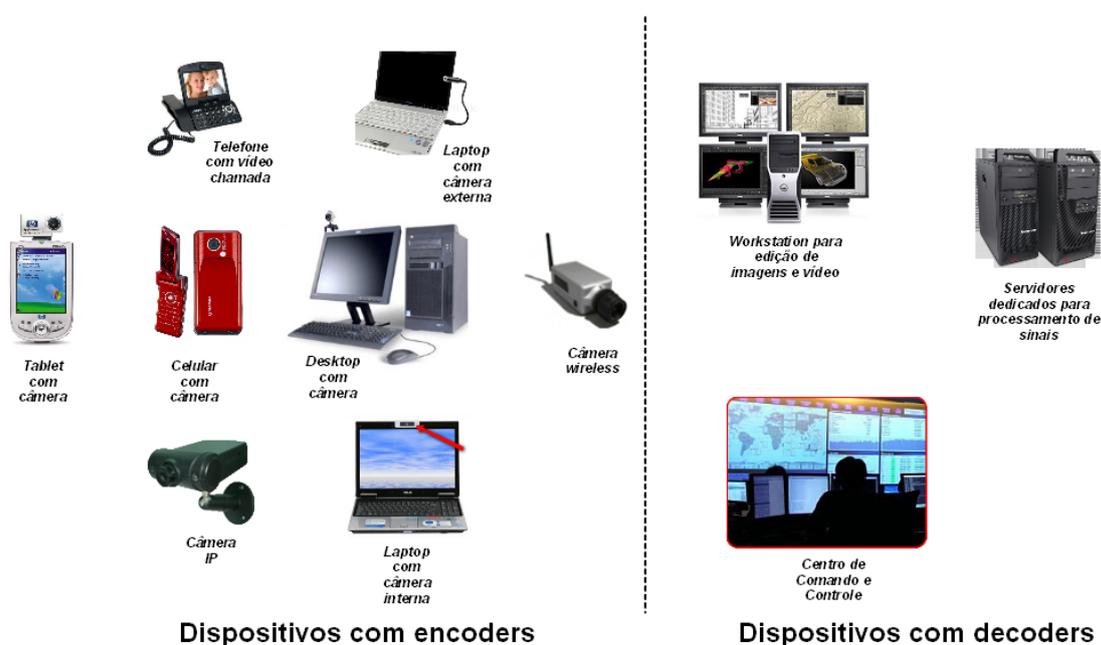


Figura 1.3 – Dispositivos utilizados na nova arquitetura

Outro ponto a ser observado é que em alguns novos cenários pode ocorrer a mudança da cardinalidade entre a fonte codificadora e o decodificador. No cenário anterior a codificação é feita apenas uma vez e a decodificação é feita n vezes, fazendo com que o esforço dispensado na codificação valha a pena, posto que será único. No novo cenário, entretanto, a codificação é feita n vezes, uma por cada dispositivo cliente e a decodificação uma única vez, num posto ou centro de processamento com melhores condições.

Essas mudanças na arquitetura dos sistemas de vídeo e nos requisitos de seus dispositivos trouxeram a necessidade de se desenvolver novos modelos de codificadores (*codecs*), uma vez que os modelos mais conhecidos não se mostraram adequados a essa nova topologia. Essa nova família de implementações recebeu o nome de **Codificação Distribuída de Vídeo** (*Distributed Video Coding* – **DVC**), que faz parte de um ramo da Teoria da Informação [1] chamado **Codificação de Fonte Distribuída** (*Distributed Source Coding* – **DSC**) [2] e será alvo de estudo dessa dissertação.

1.2

Aplicações de Codificação Distribuída de Vídeo

As novas topologias das tecnologias de telecomunicações emergentes, a mobilidade dos novos dispositivos eletrônicos e a ubiquidade das novas redes de dados foram grandes incentivadores da mudança de paradigma em relação à arquitetura dos *codecs* de vídeo.

Novos requisitos em relação a esses aspectos, conjuntamente com necessidades de autonomia e baixo consumo de energia, fizeram com que novas pesquisas fossem desenvolvidas para buscar soluções mais eficientes e que atendessem a essa demanda.

Nesse cenário, o DVC surge como solução para alguns tipos de aplicações explicadas a seguir. Umas específicas para um restrito nicho de mercado, mas com alto valor tecnológico agregado, outras mais gerais e, certamente, presentes em nosso cotidiano futuro.

1.2.1

Redes de Sensores com Suporte a Vídeo

Com o avanço das pesquisas em redes de sensores, muitas aplicações nesse segmento podem ser beneficiadas se a capacidade de monitoração visual for agregada aos elementos de rede. Isso se transforma em um diferencial quando se trata de vigilância e segurança. Essas características desejáveis devem ser agregadas, porém não podem conflitar com os requisitos básicos de uma rede, que

pode ter centenas ou milhares de elementos, todos com baixa complexidade de construção, baixo custo de produção e baixo consumo de energia.

O deslocamento da complexidade computacional do codificador para o decodificador, que é uma das principais características da arquitetura DVC, se mostra uma ótima solução para esse tipo de problema, pois os dados adquiridos pelos sensores são tratados por um aparato com alta capacidade computacional na central que consolida os dados dos elementos de rede.

Um exemplo interessante da utilização de redes de sensores com suporte a vídeo é a monitoração da vida selvagem e levantamento de parâmetros da vida animal. Um desses projetos pode ser observado em [37]. Neste cenário, a utilização do DVC parece promissora. O cenário utilizado no projeto referenciado é ilustrado na figura 1.4.

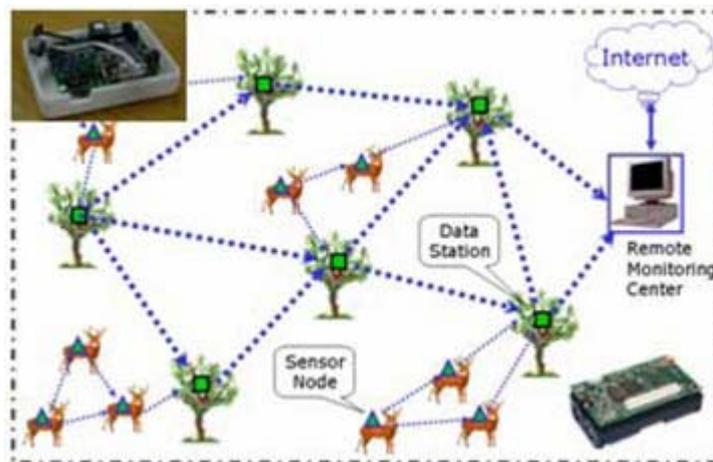


Figura 1.4 – Rede de sensores com suporte de vídeo para monitoração da vida selvagem

Fonte: Ref [37]

1.2.2

Vídeo Chamada em Serviços Móveis

Outro ramo emergente que pode aproveitar as possibilidades do DVC é o da vídeo chamada em serviços móveis. Em cada um dos terminais normalmente há um dispositivo com baixa capacidade de processamento, como um telefone celular. Assim, o sistema assimétrico com baixa complexidade no codificador

definido no DVC passa a ser interessante para quem está realizando a vídeo chamada, pois o algoritmo de codificação é de baixa complexidade [43].

O sinal, então, é enviado para uma central de transcodificação, que tem maior capacidade computacional e, após identificar o destino, pode enviar a chamada com um esquema de codificação híbrido como algum dos padrões da família MPEG-x ou H.26x, onde temos um codificador complexo, mas que vai funcionar no transcodificador. Esse envio ocorre sem qualquer problema, já que a central suporta tal complexidade. No dispositivo final do cliente, que receberá a vídeo chamada, o vídeo é decodificado com um algoritmo mais simples da mesma família MPEG-x ou H.26x, pois temos também desse lado um dispositivo eletrônico com capacidade limitada. Esse cenário de vídeo chamada com transcodificação é ilustrado na figura 1.5.

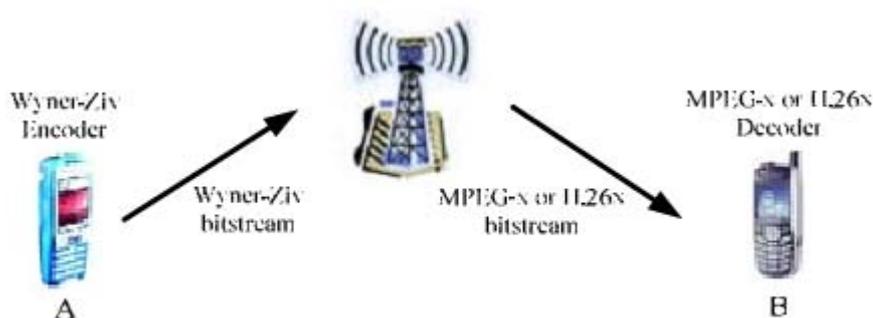


Figura 1.5 – Cenário de vídeo chamada com transcodificação, utilizando codificação Wyner-Ziv e codificação híbrida

Fonte: Ref [43]

1.2.3

Aquisição Distribuída de Vídeo

Efeitos especiais, aquisições especiais de imagens e renderização 3D podem ser feitas através de um arranjo matricial com câmeras especiais para captura de imagens. Devido à grande quantidade de informações transmitidas para uma central de processamento, uma arquitetura como a do DVC pode reduzir o custo do arranjo sem comprometer o produto final, já que a parte pesada da reconstrução é feita no computador que recebe os dados. A foto da figura 1.6 mostra a estrutura necessária para aquisição distribuída de imagem e vídeo,

conhecida como *multi-camera array*. A foto abaixo é de um dos projetos de *multi-camera array* da Universidade de Stanford [44].



Figura 1.6 – Projeto de *multi-camera array* da Universidade de Stanford

Fonte: Ref [44]

1.2.4

Rede Sem Fio de Vigilância de Baixa-Potência

O mundo digital trouxe várias vantagens tanto para usuários comuns quanto para a área governamental e de segurança. No que tange à segurança, uma evolução importante é a monitoração de áreas urbanas com sistemas de vigilância por câmeras. Para reduzir o custo de instalação e aumentar a velocidade do aumento da cobertura, muitas cidades optam pela utilização de câmeras sem fio. Porém, com o aumento da cobertura, e conseqüentemente da quantidade de câmeras, mecanismos tradicionais de codificação podem tornar o sistema caro. O custo pode aumentar ainda mais se considerarmos a vigilância em áreas rurais ou remotas, onde a falta de acesso a pontos de energia elétrica tornará necessária a utilização de câmeras com bateria [45].

No cenário acima descrito, a utilização do DVC se torna vantajosa tanto pela redução do custo de produção da câmera, diante de sua simplicidade, quanto pela redução do consumo de bateria, dando ao perfil do sistema de câmeras a qualidade de ser de baixa-potência. Tal esquema se mostra bastante razoável se considerarmos que, numa mesma área geográfica, serão necessárias várias câmeras, e a codificação será realizada por múltiplos dispositivos. Já a

decodificação, composição e otimização das imagens, serão feitas num centro de controle com mais recursos e equipamentos, obrigatoriamente mais potentes e, tão somente, uma única vez, como ilustrado na figura 1.7.

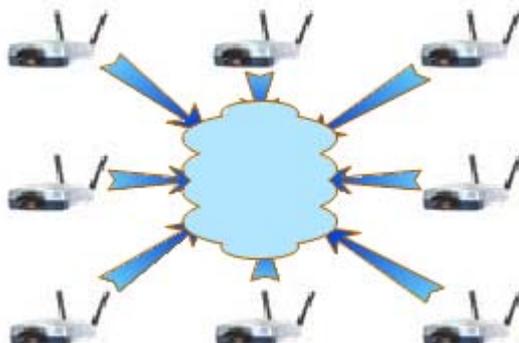


Figura 1.7 – Ilustração do cenário de Rede Sem Fio de Vigilância de Baixa-Potência

Fonte: Ref [45]

1.2.5

Câmeras Distribuídas em Veículos e Armas num Campo de Combate

Uma aplicação muito interessante no campo militar é a utilização do DVC como *codec* em Sistemas Militares de Comando e Controle (C²) [46], onde câmeras com codificadores de baixa complexidade devem ser utilizadas, já que a topologia de um campo de combate é altamente complexa considerando a diversidade e quantidade de fontes, aliadas à mobilidade difusa dos elementos dessa rede e ao requisito de baixo consumo de energia, pois o soldado não pode carregar uma quantidade grande de baterias reservas devido ao peso extra que isso acarretará em sua carga de combate.

Assim, todos os componentes da rede, tais como um soldado com sua câmera no rifle ou no capacete, carros de combate pelo terreno e aviões de vigilância enviarão imagens para serem fundidas e processadas num Centro de Comando e Controle com alto poder computacional na retaguarda dos batalhões, onde os estrategistas utilizarão as imagens das mais diversas fontes e dos mais diversos ângulos para tomarem decisões do campo de batalha.

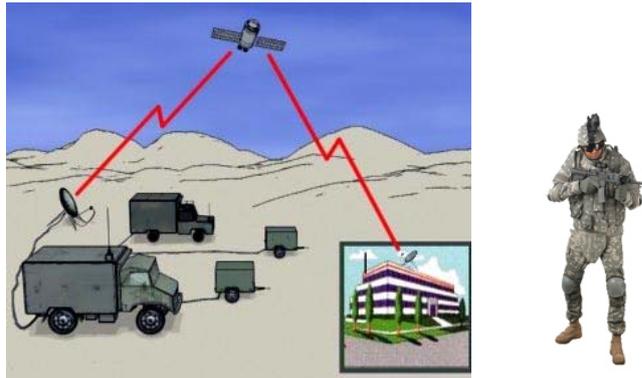


Figura 1.8 – Recursos de combate utilizados como elementos de rede num sistema de Comando e Controle

Fonte: Ref [46]

Nesta arquitetura, conforme ilustrado na figura 1.8, um soldado numa área isolada envia via rádio (na mochila) as imagens capturadas por seu capacete para a viatura de Comando e Controle, onde seu comandante de grupo analisa seus alvos e campo de atuação para tomada de decisão e distribuição das ordens. Além disso, as imagens podem ser enviadas para centros de mais alto nível na retaguarda por satélite ou enlace de microondas. Neste cenário, o codificador utilizado pelo soldado deve ter um perfil mais simplificado, contribuindo tanto na redução do seu custo de produção quanto no aumento da autonomia no campo de batalha.

A figura 1.9 mostra em (a) uma foto de um veículo de Comando e Controle da empresa BAE Systems utilizado pelo Exército Norte-Americano e em (b) uma foto de um centro de Comando e Controle à retaguarda do teatro de operações.

Vários projetos militares no mundo estão trabalhando em sistemas de vídeo, para utilização em grupos de pequenas frações com soldados. Um dos mais bem sucedidos que se tem conhecimento, pois muitos são reservados, é o projeto francês FELIN, acrônimo para *Fantassin a Equipements et Liaisons INtegres* em francês [47], ilustrado nas fotos da figura 1.10, cuja atenção, considerando a área de estudo desta dissertação, deve ser dada ao capacete balístico com sistema de vídeo integrado à cadeia de comando e controle. Outros sistemas de vídeo também são integrados ao equipamento completo como luneta infravermelho e intensificador de imagens, localizado no armamento individual do soldado (fuzil francês FAMAS 5.56 mm).



(a)



(b)

Figura 1.9 – Em (a) foto da viatura de Comando e Controle da empresa BAE Systems utilizado pelo Exército Norte-Americano e em (b) foto de centro de Comando e Controle
 Fonte: BAE Systems, disponível em <<http://www.baesystems.com>>, acesso em 15 de setembro de 2010

No Exército Brasileiro, projeto similar está na fase de levantamento de requisitos. O chamado projeto COBRA (Combatente Brasileiro do Futuro) tem previsão de lançamento dos primeiros protótipos a partir de 2016.



(a)



(b)

Figura 1.10 – Detalhes do sistema captura de vídeo do projeto francês FELIN
 Fonte: Ref [47]

Na figura 1.10, temos em (a) o detalhe do capacete com sistema de vídeo e em (b) da luneta no armamento. Vale destacar também, em (a), o cabo que sai do

capacete ligando a câmera ao sistema de baterias e rádio junto ao corpo do soldado.

Na figura 1.11, temos duas fotos onde vemos a utilização do sistema de vídeo da luneta no armamento do soldado, auxiliando no estudo e visualização segura do campo de visão com o auxílio de um *tablet* que recebe as informações. Em (a) temos a solução do Exército Francês e em (b) o sistema utilizado pelo Exército Australiano.



Figura 1.11 – Utilização do sistema de vídeo e transmissão/recepção de dados. Em (a) temos o projeto do Exército Francês e em (b) a solução utilizada pelo Exército Australiano

Fonte: Ref [47]

1.3

Proposta da Dissertação

Dos pontos de vista prático e de viabilidade econômica, o DVC ainda fomenta estudos e pesquisas com o intuito de amadurecer as arquiteturas candidatas e seus algoritmos propostos a fim de torná-los uma realidade para as aplicações alvo [48].

Procurando conhecer a tecnologia e contribuir com os estudos acima citados, esta dissertação seguirá a seguinte metodologia de estudo para atingir os objetivos que serão apresentados (em ordem cronológica dos trabalhos realizados):

- Estudar os fundamentos e técnicas de codificação de vídeo tradicionais;
- Estudar a Teoria da Informação de Shannon [1] e as teorias correlatas aplicadas à Codificação de Fonte Distribuída (*Distributed Source Coding* – DSC) [2] e Codificação Distribuída de Vídeo (*Distributed Video Coding* – DVC) [3, 4];
- Estudar as teorias dos blocos e dos algoritmos aplicados às arquiteturas clássicas de DVC, disponíveis na literatura acadêmica;
- Levantar as possíveis aplicações alvo para utilização do DVC.

A partir da consolidação dos estudos acima apresentados, pretende-se atingir os seguintes objetivos:

- Propor uma arquitetura candidata de DVC;
- Implementar um *framework* computacional para a arquitetura proposta em linguagem interpretada;
- Construir um projeto orientado a objetos da arquitetura proposta;
- Implementar o projeto orientado a objetos numa linguagem coerente com este paradigma e de alto desempenho;
- Realizar experimentos e coletar resultados com as implementações de DVC construídas;
- Propor trabalhos futuros de melhoria e continuidade a partir das conclusões observadas.

1.4

Organização da Dissertação

Esta dissertação está dividida em **sete capítulos**, incluindo o presente capítulo introdutório, o **Capítulo 1**, mais **dois apêndices** complementares.

Como ambientação, o **Apêndice A** aborda alguns conceitos e princípios básicos de processamento de vídeo, necessários para o entendimento desta dissertação.

Como premissa, este trabalho apresenta a teoria matemática que abriu a possibilidade para se alcançar o desempenho alvo e o projeto desta nova arquitetura de codificação de vídeo, o DVC, o que é feito no **Apêndice B**, através da apresentação da Teoria da Informação de Shannon e dos Teoremas de Slepian-Wolf e Wyner-Ziv, que são as bases para o entendimento da Codificação de Fonte Distribuída.

As arquiteturas clássicas que implementaram este novo paradigma são apresentadas no **Capítulo 2**, onde são descritas as três arquiteturas consideradas o “estado da arte” em DVC, uma proposta pelo grupo de estudo da Universidade de Stanford, modelo DVC de Stanford [25, 26, 27, 28, 29], outra pelo grupo de estudo da Universidade de Berkeley, com a implementação conhecida como PRISM [30, 31, 32] e a última, que produziu e publicou os resultados mais completos e recentes, originada de um consórcio de universidades europeias, que desenvolveram o projeto DISCOVER [48].

No **Capítulo 3**, propõem-se modificações aos esquemas clássicos de DVC, dando origem a uma nova Arquitetura Conceitual, que chamamos de **Open DVC**, a qual é implementada neste trabalho numa plataforma aberta.

No **Capítulo 4**, apresentamos a topologia e organização de um framework aberto orientado a objetos para a implementação da Arquitetura Conceitual proposta, ou seja, sua Arquitetura Computacional.

No **Capítulo 5**, são apresentados os experimentos realizados com sequências padrão de vídeo bruto nas simulações e seus respectivos resultados.

As conclusões e as sugestões para trabalhos futuros são abordadas no **Capítulo 6**, onde são também detalhadas as contribuições feitas, as publicações e as perspectivas para a continuidade da pesquisa.

Por último, no **Capítulo 7** são apresentadas as referências bibliográficas utilizadas nesta dissertação.