

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
CIÊNCIA DA COMPUTAÇÃO

André Filipe da Silva Fernandes

Solução para o problema da taxa alvo durante a codificação de light fields

Florianópolis
23 de novembro de 2022

André Filipe da Silva Fernandes

Solução para o problema da taxa alvo durante a codificação de light fields

Trabalho de Conclusão de Curso submetido ao Curso de Graduação em Ciência da Computação do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Ismael Seidel, Dr.

Coorientador: Prof. José Luís A. Güntzel (?), Dr.

Florianópolis
23 de novembro de 2022

RESUMO

Em algoritmos de compressão de dados com perdas existe uma necessidade natural de balanceamento da taxa de compressão com a taxa de distorção, pois taxas de compressão muito elevadas produzem arquivos de má qualidade, enquanto taxas de compressão muito baixas proporcionam arquivos de tamanhos impraticáveis. A parte 2 do padrão JPEG Pleno, que é dedicada à codificação de lightfields, possui um modo de transformadas 4D que utiliza um multiplicador Lagrangiano para realizar este balanceamento (ALVES et al., 2020). O problema de se utilizar o multiplicador Lagrangiano para este fim é a dificuldade em estimar o tamanho dos dados de saída, e portanto, escolher os parâmetros de maneira adequada. Este trabalho propõe um algoritmo capaz de otimizar a taxa de distorção baseando-se em uma taxa alvo de bits por píxel, e a implementação deste algoritmo na parte 2 do software de referência do padrão JPEG Pleno para o modo de transformadas 4D.

Palavras-chave: Lightfields. JPEG Pleno. Compressão de dados.

SUMÁRIO

1	INTRODUÇÃO	7
1.1	OBJETIVOS	8
1.1.1	Objetivo Geral	8
1.1.2	Objetivos Específicos	8
1.2	MÉTODO DE PESQUISA	8
1.3	CRONOGRAMA	8
1.4	CUSTOS	8
1.5	RECURSOS HUMANOS	8
1.6	COMUNICAÇÃO	8
1.7	RISCOS	8
	REFERÊNCIAS	9

1 INTRODUÇÃO

Light fields são uma modalidade de representação de imagens e vídeos que tenta expandir a capacidade representativa destes meios através da captura de múltiplos ângulos simultaneamente. Através disso, os light fields proporcionam funcionalidades novas perante o paradigma já estabelecido, como simular o efeito de paralax através de tipos especiais de telas, permitindo uma experiência mais natural e imersiva do que no cinema 3D tradicional. Também é possível transformar light fields em imagens tradicionais com a vantagem de poder realizar pequenas modificações de ângulo ou ajustar o foco nas etapas de pós processamento. Light fields podem ser capturados através de um agrupamento de câmeras, ou através de arrays de microlentes.

Assim como em outras modalidades de imagens, light fields em sua forma bruta exigem uma quantidade impraticavelmente grande de memória para armazenamento, especialmente devido à natureza quadridimensional destes dados que codificam tanto a posição quanto o ângulo incidente em cada ponto da imagem, portanto a utilização de técnicas de compressão são desejáveis. Neste contexto foi criado o grupo JPEG Pleno com a intenção de estabelecer padrões para codificação e representação de light fields e de outras imagens plenópticas, como point clouds e hologramas. Até o momento o padrão inclui dois modos para codificar light fields: 4D-Transform mode, que é baseado na transformada discreta de cosseno em 4 dimensões - *4 dimensional discrete cosine transform* (4D-DCT); e 4D-Prediction mode que é preditivo, capaz de prever estágios intermediários de partes da imagem baseado no contexto.

O modo de transformada 4D, no qual este trabalho se concentra, particiona a imagem em blocos quadridimensionais não sobrepostos, limitados por um tamanho máximo configurável. Com estes blocos criados, é possível dividi-los recursivamente em até 16 blocos menores, seguindo algum critério de taxa de distorção (preferencialmente concentrando mais blocos em partes mais detalhadas da imagem) de forma que o custo Lagrangiano seja minimizado, conforme definido na equação 1.1. A forma como estes blocos foram divididos é modelada como uma hexadecatree e codificada em um bitplane, para possibilitar a reconstrução posterior da imagem, visto que eles passarão por uma 4D-DCT.

Ter um controle do tamanho dos dados após comprimidos é uma característica útil para codecs, especialmente se utilizado em aplicações de streaming, por exemplo, onde é necessário controlar precisamente a vazão dos dados de acordo com a capacidade da rede. A utilização de um custo Lagrangiano para o particionamento dos blocos torna difícil a tarefa de estimar a taxa alvo através dos parâmetros de entrada. O presente trabalho resolve este problema modificando o algoritmo de particionamento para que ele se limite a um tamanho fixo de saída que será passado como parâmetro de entrada, dispensando o uso do operador Lagrangiano.

$$J = D + \lambda R \quad (1.1)$$

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O objetivo deste trabalho é implementar no software de referência do JPEG Pleno um algoritmo capaz de otimizar a taxa de distorção baseando-se em uma taxa alvo de bits por píxel.

1.1.2 Objetivos Específicos

- Demonstrar que o algoritmo proposto atinge a taxa alvo conforme o esperado.
- Avaliar se os light fields gerados através da taxa alvo mantêm qualidade semelhante àsqueles gerados através do multiplicador Lagrangiano.
- Avaliar a eficiência e desempenho do novo algoritmo em relação ao anterior.

1.2 MÉTODO DE PESQUISA

No início do projeto será feito o estudo da base de código do software de referência do JPEG Pleno, bem como a leitura de trabalhos correlatos ao tema proposto, visando conhecer técnicas empregadas especialmente em codecs de imagem e vídeo e que podem ser utilizadas no desenvolvimento deste trabalho. A partir do conhecimento adquirido através deste estudo, o algoritmo será desenvolvido a partir do software de referência, utilizando a linguagem c++. Assim que o algoritmo estiver funcional serão feitas as comparações de qualidade e desempenho com o modelo original, utilizando os lightfields fornecidos pelo dataset de light fields do JPEG Pleno com parâmetros variados. As comparações de qualidade serão feitas através da Relação Sinal-Ruído de Pico - *Peak signal to noise ratio* (PSNR), enquanto o desempenho será medido em tempo de execução.

1.3 CRONOGRAMA

1.4 CUSTOS

1.5 RECURSOS HUMANOS

1.6 COMUNICAÇÃO

1.7 RISCOS

REFERÊNCIAS

ALVES, G. D. O. et al. The jpeg pleno light field coding standard 4d-transform mode: How to design an efficient 4d-native codec. **IEEE Access**, v. 8, p. 170807–170829, 2020.

PERRA, C. et al. An overview of the emerging JPEG Pleno standard, conformance testing and reference software. In: SCHELKENS, P.; KOZACKI, T. (Ed.). **Optics, Photonics and Digital Technologies for Imaging Applications VI**. SPIE, 2020. v. 11353, p. 207 – 219. Disponível em: <https://doi.org/10.1117/12.2555841>.