

O Papel das Primeiras Técnicas de Fusão:

Eduardo Kenji Avena

Universidade de São Paulo - USP, São Paulo-SP, Brasil

edukenjister@gmail.com

RESUMO

As primeiras técnicas de fusão foram propostas em tarefas de análise de vídeo como uma forma de melhorar a eficácia, eficiência e valores, gerando modelos de dados compactos capazes de manter pistas semânticas presentes em dados multimodais. As primeiras tentativas de fundir dados multimodais empregaram operadores de fusão no espaço de recursos de baixo nível, perdendo a representatividade dos dados. Isso levou a esforços de pesquisa posteriores para evoluir operadores simples para operações complexas, que se tornaram, em geral, inseparáveis do processamento de pistas semânticas multimodais. Neste artigo, investigamos a aplicação dos primeiros operadores de fusão multimodal no espaço de recursos de nível médio. Cinco operadores diferentes (concatenação, soma, grama, média e máximo) foram empregados para fundir recursos de vídeo multimodal de nível médio. Os dados fundidos derivados de cada operador foram então usados como entrada para duas tarefas de análise de vídeos diferentes: Segmentação de Cena de Vídeo Temporal e Classificação de Vídeo. Para cada tarefa, realizamos uma análise comparativa entre os operadores e técnicas de trabalho relacionadas projetadas para essas tarefas usando operações de fusão complexas. Os resultados de eficácia alcançados pelos operadores foram muito próximos aos alcançados pelas técnicas, apontando fortes evidências de que trabalhar em um espaço de recursos mais homogêneo pode reduzir desvantagens conhecidas de fusão de baixo nível. Além disso, os operadores tornam a fusão de dados separável, permitindo que os pesquisadores mantenham o foco no desenvolvimento de representações de pistas semânticas.

Referências

A. E. Abdel-Hakim e A. A. Farag. 2006. CSIFT: A SIFT Descriptor with Color Invariant Characteristics. Em 2006, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 2. 1978-1983. Biblioteca Digital do Google Scholar

David Arthur e Sergei Vassilvitskii. 2006. k-means ++: The Advantages of Careful Seeding. Relatório técnico 2006--13. Stanford InfoLab. <http://ilpubs.stanford.edu:8090/778/Google Scholar>

Lorenzo Baraldi, Costantino Grana e Rita Cucchiara. 2015. A Deep Siamese Network for Scene Detection in Broadcast Videos. In Proceedings of the 23rd ACM International Conference on Multimedia (MM '15). Association for Computing Machinery, New York, NY, USA, 1199--1202. <https://doi.org/10.1145/2733373.2806316> Google ScholarDigital Library

Mostrar todas as referências

Termos do Índice

Avaliando os primeiros operadores de fusão no espaço de recursos de nível médio

Metodologias de computação

Inteligência artificial

Visão computacional

Problemas de visão computacional

Segmentação de vídeo

Sistemas de informação

Recuperação de informação

Modelos de recuperação e classificação

Combinação, fusão e pesquisa federada

Tarefas e objetivos de recuperação

Extração de informação

Aplicativos de sistemas de informação

Sistemas de informação multimídia

Perguntas:

- 1)"Qual a principal mensagem do trabalho?" Para responder a esta pergunta, identifique no artigo qual o problema que os autores procuram resolver
- 2)"Qual a principal contribuição do trabalho?" Para responder a esta pergunta, identifique no artigo qual a solução reportada pelos autores
- 3)"Como o trabalho foi avaliado?" Para responder a esta pergunta, identifique no artigo como os autores avaliaram o impacto da solução proposta no problema identificado originariamente
- 4)"Qual a principal limitação do trabalho?" Para responder esta pergunta, *analise criticamente* o trabalho apresentado no artigo e identifique uma limitação na solução reportada relativamente ao problema tratado

Sua resposta

- 1) Perda de representatividade de dados computacionais, modelos de dados computacionais extensos e que ocupam muito espaço em bytes de armazenamento em Hard Disk (Disco Rígido) ou Cloud Computing (Computação em Nuvens), perdas de pistas semânticas presentes em dados computacionais multimodais;
- 2) As primeiras técnicas de fusão ou junção foram propostas em tarefas de análise de vídeo como uma forma de melhorar a eficiência, eficácia e valores, gerando modelos de dados computacionais compactos e enxutos capazes de manter pistas semânticas presentes em dados computacionais multimodais. As primeiras tentativas de fundir dados computacionais multimodais empregaram operadores de fusão no espaço de recursos de baixo nível, perdendo a representatividade dos dados computacionais. Isso levou a esforços de pesquisas posteriores para evoluir operadores simples para operações complexas, que se tornaram, em geral, inseparáveis do processamento de pistas semânticas multimodais. Neste artigo, investigamos a aplicação dos primeiros operadores de fusão computacional multimodal no espaço de recursos de nível médio. Cinco operadores diferentes (concatenação, soma, grama, média e máximo) foram empregados para fundir recursos de vídeo multimodal de nível médio. Os dados computacionais fundidos derivados de cada operador foram então usados como entrada para duas tarefas de análise de vídeos diferentes: Segmentação de Cena de Vídeo Temporal e Classificação de Vídeo. Para cada tarefa, os autores realizaram uma análise comparativa entre os operadores e técnicas de trabalho relacionadas projetadas para essas tarefas usando operações de fusão complexas;
- 3) Os resultados de eficiência e eficácia alcançados pelos operadores de fusão de dados computacionais foram muito próximos aos alcançados pelas técnicas de fusão ou junção, apontando fortes evidências de que trabalhar em um espaço de recursos mais homogêneo e uniforme pode reduzir desvantagens conhecidas de fusão de baixo nível. Além disso, os operadores tornam a fusão ou junção de dados computacionais separável, permitindo que os pesquisadores mantenham o foco no desenvolvimento de representações de pistas semânticas;
- 4) Limitação do trabalho quanto aos recursos financeiros, de mão-de-obra e materiais (suprimentos), destinados a pesquisa e desenvolvimento de técnicas de fusão de dados computacionais enxutos e compactos.