Data in Brief 51 (2023) 109627

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

CineScale2: a dataset of cinematic camera features in movies

Mattia Savardi[a], András Bálint Kovács[b], Alberto Signoroni[a], Sergio Benini[c,*]

[a] Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health – University of Brescia, Brescia, Italy
[b] Film Department, ELTE University, Budapest, Hungary
[c] Department of Information Engineering, University of Brescia, Brescia, Italy

ARTICLE INFO

Article history:
Received 30 May 2023
Revised 9 August 2023
Accepted 22 September 2023
Available online 28 September 2023

Dataset link: CineScale2 (Original data)

Keywords:
Camera features
Camera angle
Camera level
Movie frames
Video analysis
Movies
Film studies

ABSTRACT

The position and orientation of the camera in relation to the subject(s) in a movie scene, namely *camera "level"* and *camera "angle"*, are essential features in the film-making process due to their influence on the viewer's perception of the scene. We provide a database containing camera feature annotations on camera angle and camera level, for about 25,000 image frames. Frames are sampled from a wide range of movies, freely available images, and shots from cinematographic websites, and are annotated on the following five categories - Overhead, High, Neutral, Low, and Dutch - for what concerns camera angle, and on six different classes of camera level: Aerial, Eye, Shoulder, Hip, Knee, and Ground level. This dataset is an extension of the Cinescale dataset [1], which contains movie frames and related annotations regarding shot scale. The CineScale2 database enables AI-driven interpretation of shot scale data and opens to a large set of research activities related to the automatic visual analysis of cinematic material, such as movie stylistic analysis, video recommendation, and media psychology. To these purposes, we also provide the model and the code for building a Convolutional Neural Network (CNN) architecture for automated

* Corresponding author.
E-mail address: sergio.benini@unibs.it (S. Benini).

https://doi.org/10.1016/j.dib.2023.109627

camera feature recognition. All the material is provided on the the project website; video frames can be also provided upon requests to authors, for research purposes under fair use.

Specifications Table

Subject	Arts (General), Film studies, Video content analysis
Specific subject area	Quantitative analysis of camera angle and camera level in films
Type of data	Images (.JPG format)
How data were acquired	Camera angle and camera level annotations were provided by three human coders (2 coders + 1 who made decision in case of disagreement) in tabular form.
Data format	Raw
Description of data collection	The database contains the camera angle and camera level annotations for about 25,000 frames. Camera angle is labelled on five categories - Overhead, High, Neutral, Low, and Dutch - while camera level on six classes: Aerial, Eye, Shoulder, Hip, Knee, and Ground level. In total, 24,748 frames were sampled from a wide range of movies, freely available images scraped online, and shots from cinematographic websites; 9129 frames were annotated for camera angle, while remaining 15,619 for camera level.
Data source location	Institution: University of Brescia City/Town/Region: Brescia Country: Italy Latitude: 45.564664 Longitude: 10.231660
Data accessibility	Repository name: Mendeley Data Data identification number: 10.17632/h4n3gn93gz.3 Direct URL to data: https://data.mendeley.com/datasets/h4n3gn93gz/3 Project website: https://cinescale.github.io

1. Value of the Data

- The data can be used by both computer scientists and film scholars to perform quantitative analysis of camera features in movies.
- The data are useful to develop models and classification strategies to automatically predict the camera features in movies.
- There are evidences that the statistical distribution and the temporal pattern of some frequent cinematic features might act as a stylistic fingerprint of a specific director. By evaluating the statistics of such cinematographic features, it is possible to investigate of how camera features influence the film viewing experience in viewers, such as empathy related processes.

2. Objective

The position and orientation of the camera in relation to the subject(s) in a movie scene, namely camera “level” and camera “angle”, are essential features in the film-making process due to their influence on the viewer's perception of the scene. We provide a database [11] containing camera feature annotations on camera angle and camera level, for more than 25,000 image frames that can be helpful for a large set of research activities related to the automatic visual analysis of cinematic material, such as movie stylistic analysis, video recommendation, and media psychology.

3. Data Description

Data consists of camera angle and camera level annotations for a total of 24,748 frames. Specifically, 9129 video frames are categorized according to their camera angle in the following five categories: Overhead, High, Neutral, Low, and Dutch angle. The remaining 15,619 video frames are labelled into six different classes accounting for their camera level: Aerial, Eye, Shoulder, Hip, Knee, and Ground. The exact numbers of frames in datasets and classes of camera angle and level are illustrated in Table 1.

Table 1
Number of frames for each class in the dataset.

ANGLE	Overhead	High	Neutral	Low	Dutch	
	470	4956	687	2031	985	
LEVEL	Aerial	Eye	Shoulder	Hip	Knee	Ground
	7152	3513	3006	1332	249	367

3.1. Categories

In Figure 1, we provide some visual examples which illustrate how camera angle is mapped into the six categories, and a reference scheme for different angles. The five classes shown in

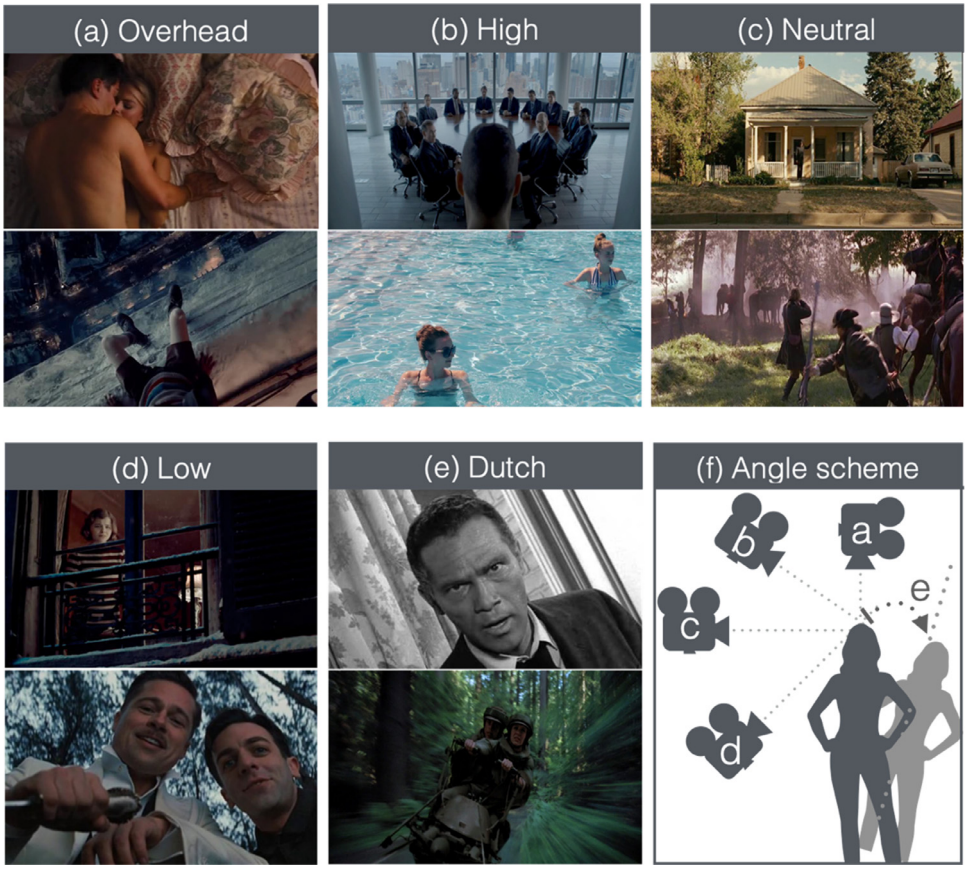

Fig. 1. Examples of different camera angle classes (a-e), and a reference scheme (f).

Figure 1 (i.e., Overhead, High, Neutral, Low, and Dutch) describe camera rotation along both lateral (High, Neutral, and Low) and longitudinal (Dutch) axes. In particular, an Overhead-angle indicates a take looking down on a subject from an almost perpendicular direction.

Figure 2 instead provides a reference scheme and examples of different camera levels. In particular, we show how camera level (i.e., the height of the camera in the scene in relation to the subject being framed) is categorized into six different classes: Aerial, Eye, Shoulder, Hip, Knee, and Ground level. The particular class of Aerial-level is used for shots taken from a considerable height, such as from a plane or a drone, showing a large portion of the surroundings.

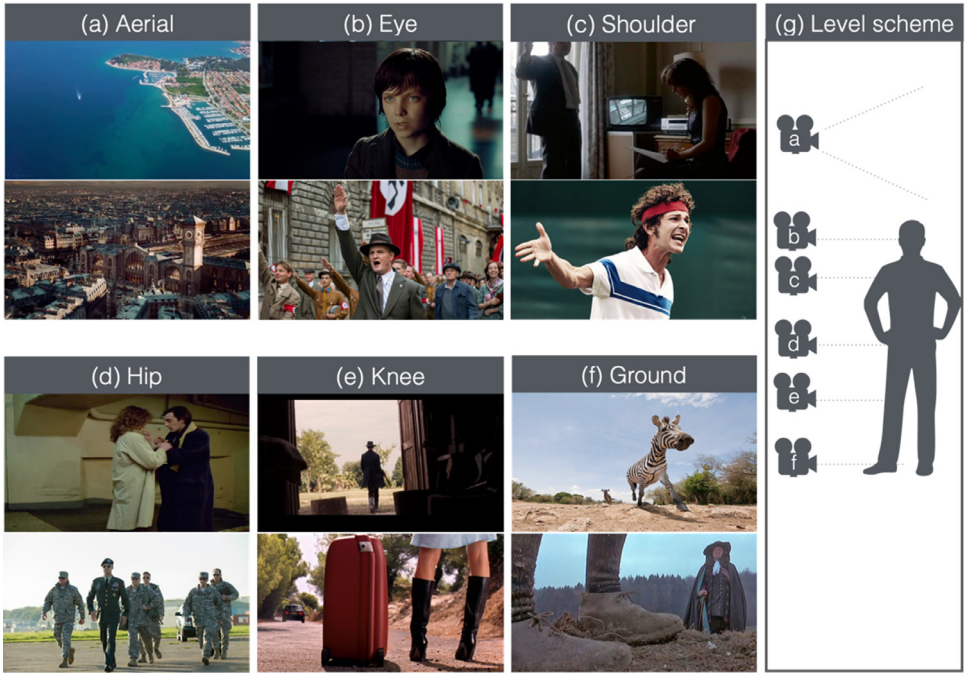

Fig. 2. Examples of different camera level classes (a-f), and a reference scheme (g).

3.2. Importance of the features

Camera angle and level are fundamental aspects to consider during movie production, alongside other critical elements such as shot scale, camera movement, and shot editing, in order to enhance the storytelling experience for viewers [2,3]. Especially in the practice of art cinema, the preferences in the use of specific camera features can be important indicators of a particular style: previous work [4] indicates for example that the statistical analysis of the overall shot scale distribution and transitions in films may reveal consistent and recurrent patterns in the works of a specific author. In [2], it is shown how this can even lead to performing automatic attribution of the movie's authorship starting from the statistical analysis of shot scale and other camera features.

The importance of camera angle lies in its ability to establish a power dynamic between characters, as prior research has demonstrated that low-angle shots (where the viewer is forced to look up at the characters) can convey dominance, strength, and aggression, while high-angle shots (where the viewer looks down at the characters) can imply weakness and vulnerability

[5]. Camera angle can also influence empathic engagement by affecting the audience's attitudes towards and evaluations of characters [6], products [7] and the credibility of a speaker in promotion videos [8,9].

Camera level, on the other hand, is a tool for controlling storytelling by determining the viewer's perspective on the scene. Eye-level shots are considered neutral and are often used to show natural conversations between characters, while knee- and ground-level takes can be used to feature characters walking without revealing their face, inducing viewers to imagine what is happening at higher levels. Aerial-level shots, in turn, provide viewers with a reference in space, time, or reality. Using different camera levels can affect the viewer's empathy, with eye-level shots promoting perceived similarity to the subject of the camera [5].

4. Experimental Design, Materials and Methods

4.1. Data extraction

In preparation for annotation, individual frames were extracted from various sources, including web scraping, online movie websites, the CineScale dataset [1], and the MovieShot dataset [10]. This diverse selection of sources ensured a comprehensive representation of movie frames for the subsequent analysis.

To eliminate duplicates or near-duplicate frames, a post-processing pipeline was implemented. This pipeline extracted color histograms and morphological features from each frame, allowing for the identification and removal of highly similar frames. Consequently, only unique frames proceeded to the annotation phase, ensuring the integrity and diversity of the dataset.

Shot annotations for camera angle and level were performed by a team of three human coders. Two primary coders were responsible for the initial annotations, while a third coder served as a decision-maker in cases of disagreement. This collaborative approach ensured the accuracy and consistency of the annotations.

Given the rarity of Dutch shots in filmographies, a synthetic augmentation block was implemented to enhance the dataset. This block randomly selected frames from the original raw data and applied random rotations. Subsequently, a human coder manually validated the augmented frames. These additional shots were incorporated into the dataset, with filenames prefixed by the keyword "gen " to indicate their origin. The annotated dataset can be accessed through the Mendeley Data repository . The dataset is organized into two primary folders, "angle" and "level", representing the two main formal features. Within each folder, frames are further categorized into subfolders according to their respective classes. This hierarchical structure facilitates efficient navigation and retrieval of specific frames for analysis in scientific publications.

Additional details and materials are available at the project website, where it possible to download a CNN model that performs camera angle and level classification, and the related performance obtained on the test set. In particular, the CNN architecture for classification is based on ResNet [12] family as the backend, and includes two independent classification heads, each responsible for categorizing a different non-exclusive set of classes.

Ethics Statement

No human subjects were involved in this work.

Data availability

CineScale2 (Original data) (Mendeley Data).

CRediT Author Statement

Mattia Savardi: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization; **András Bálint Kovács:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation; **Alberto Signoroni:** Methodology, Investigation, Resources, Writing – review & editing, Supervision, Funding acquisition; **Sergio Benini:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Acknowledgements

The author would like to thank MSc students Andrea Ferrari and Gianluca Gualandris for their contribution on camera feature analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Savardi, A.B. Kovacs, A. Signoroni, S. Benini, Cinescale: A dataset of cinematic shot scale in movies, Data´in Brief 36 (2021) 107002.
[2] M. Svanera, M. Savardi, A. Signoroni, A.B. Kovacs, S. Benini, Who is the film's director? authorship recognition´based on shot features, IEEE MultiMedia 26 (2019) 43–54.
[3] S. Benini, M. Savardi, K. Balint, A.B. Kovacs, A. Signoroni, On the influence of shot scale on film mood and narrative engagement in film viewers, IEEE Trans. Affect. Comput. 13 (2019) 592–603.
[4] G. Raz, G. Valente, M. Svanera, S. Benini, A.B. Kovacs, A robust neural fingerprint of cinematic shot-scale, Projections 13 (2019) 23–52 01 Dec..
[5] W. Huang, J.S. Olson, G.M. Olson, Camera angle affects dominance in video-mediated communication, in: CHI '02 Extended Abstracts on Human Factors in Computing Systems, CHI EA '02, Association for Computing Machinery, New York, NY, USA, 2002, pp. 716–717, doi:10.1145/506443.506562. URL.
[6] T. Lankhuizen, K.E. Balint, M. Savardi, E.A. Konijn, A. Bartsch, S. Benini, Shaping film: A quantitative formal´analysis of contemporary empathy-eliciting hollywood cinema, Psychology of Aesthetics, Creativity, and the Arts (2020).
[7] J. Meyers-Levy, L.A. Peracchio, Getting an angle in advertising: The effect of camera angle on product evaluations, J. Market. Res. 29 (1992) 454–461.
[8] R.K. Tiemens, Some relationships of camera angle to communicator credibility, J. Broadcast. 14 (1970) 483–490.
[9] T.A. McCain, J. Chilberg, J. Wakshlag, The effect of camera angle on source credibility and attraction, J. Broadcast. 21 (1977) 35–46.
[10] A. Rao, J. Wang, L. Xu, X. Jiang, Q. Huang, B. Zhou, D. Lin, A unified framework for shot type classification based on subject centric lens, in: European Conference on Computer Vision, Springer, 2020, pp. 17–34.
[11] M. Savardi, S. Benini, CineScale2 dataset, Mendeley Data (2023) V1, doi:10.17632/h4n3gn93gz.3.
[12] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.