

Cross-Camera Human Motion Transfer by Time Series Analysis

Anonymous Submission

Submitted to IEEE ICASSP'2024

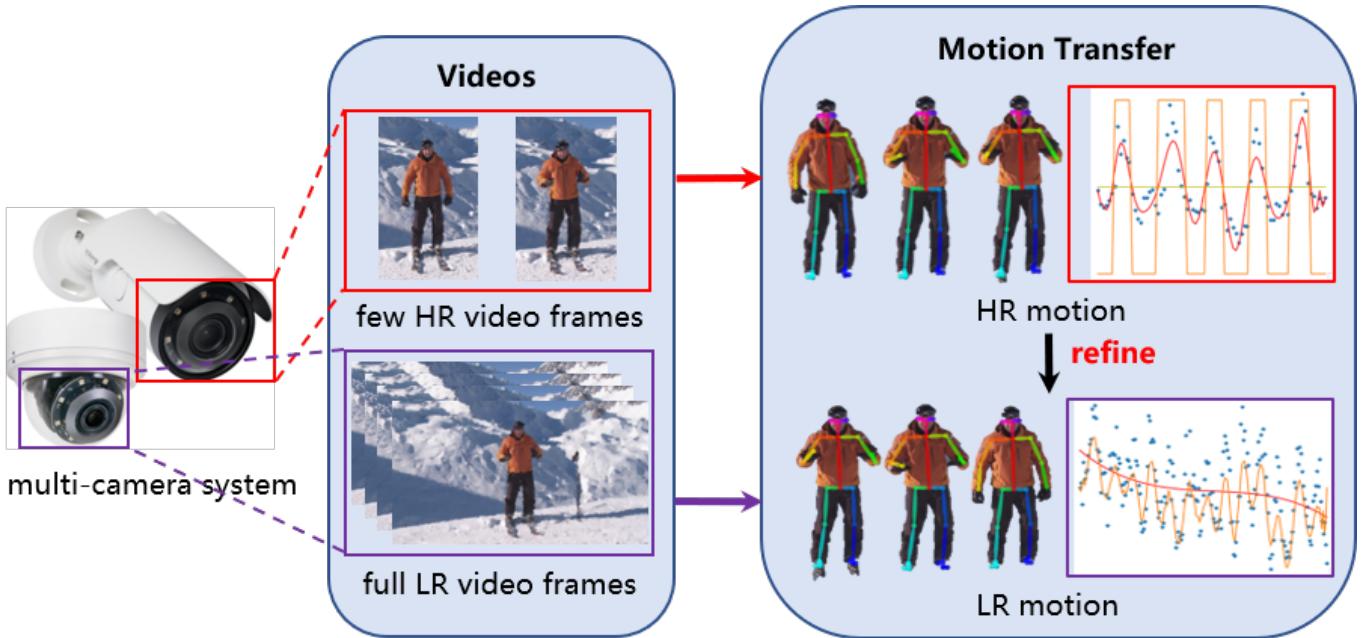


Fig. 1: In our proposed framework, we adopt a multi-camera system to capture multi-scale human-centric videos. We first estimate 3D human poses from both high-resolution (HR) and low-resolution (LR) videos. Following this, we extract human motion patterns from the HR videos. These HR patterns are then leveraged across the various camera feeds, serving to significantly enhance the quality of pose estimation initially derived from LR videos. This comprehensive methodology ensures superior pose estimation, irrespective of the video resolution.

Abstract—With the rapid advancement in optical sensor technology, the development and utilization of imaging systems that integrate heterogeneous cameras have become increasingly common, enhancing high-resolution (HR) video acquisition and analysis. Despite this progress, the intrinsic characteristics of multiple cameras create substantial challenges in motion transfer. In this work, we propose an algorithm using time series analysis to address these challenges and facilitate efficient motion transfer across multiple cameras. The process involves identifying seasonality in motion data, then constructing an additive time series model to distill transferable patterns across disparate camera feeds. Validated on real-world data, our method showcases efficiency and interpretability, courtesy of its precise mathematical formulation. Our motion transfer algorithm enhances practical utility by integrating seamlessly with downstream tasks, notably improving pose estimation in low-resolution videos through the exploitation of patterns derived from HR counterparts.

Index Terms—Motion Transfer, Time Series Analysis, Camera Systems, Computational Imaging

I. INTRODUCTION

The rapid advancements in optical sensor technology have propelled multi-camera systems to the forefront of computational imaging solutions. Such systems capitalize on a combination of short-focus and long-focus lenses, capturing expansive field-of-view (FoV) videos and high-resolution (HR) local-view details, respectively. By infusing the high-resolution details into low-resolution (LR) videos, multi-camera systems can synthesize videos that balance breadth of perspective with fine detail [1]–[3].

Given the proliferation of multi-camera applications in everyday scenarios such as smart cities and live sports, capturing multi-scale human-centric videos via hybrid-camera systems has become a prevalent trend. Consequently, there has been a surge in demand for incorporating detailed human-centric information into broader view videos. However, a significant challenge arises, particularly when dealing with motion: the

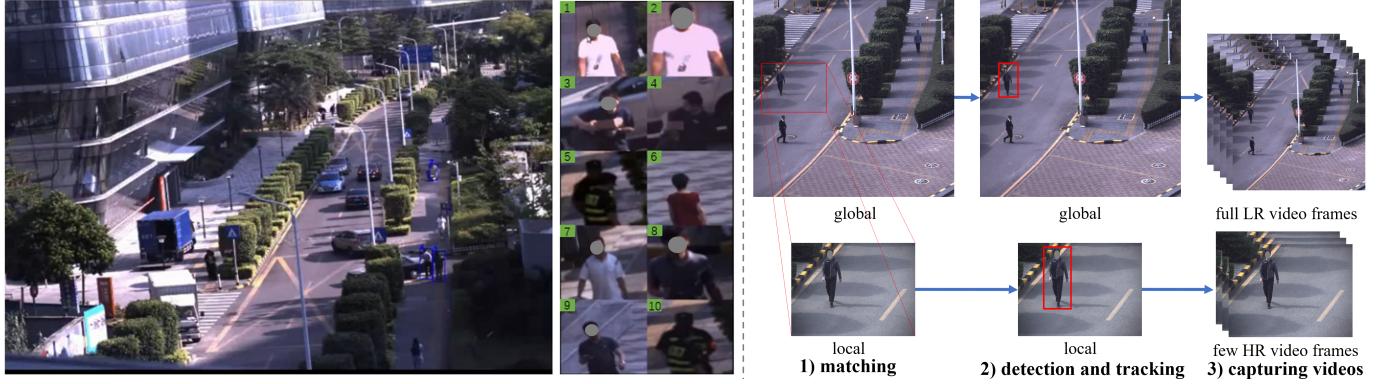


Fig. 2: A case for surveillance and security, where a multi-camera system can observe pedestrian trajectories in a large FoV video. Meanwhile, the HR details of pedestrians are captured by local-view cameras. With such a setting, it is desirable to fuse the information from both global and local views, such that interesting or suspicious objects can be recognized, and the corresponding HR human details can be extracted. Since we collect and use real-life data in this figure, the human faces are masked out to protect personal privacy.

task of effectively transferring human movement data across heterogeneous camera systems. In this paper, we delve into this intriguing and complex problem, as illustrated in Figure 1.

Figure 2 provides a glimpse into a surveillance and security setup, where a multi-camera system tracks pedestrian movements across a large FoV video while simultaneously capturing high-resolution details using local-view cameras. In such a scenario, the fusion of information from global and local views becomes desirable, enabling the identification of objects of interest and extraction of corresponding high-resolution human details. However, the multi-scale characteristics of hybrid cameras, particularly with regard to local spatial resolution, pose significant challenges to the fusion process.

Moreover, integrating multi-scale human-centric videos into a single composite video that maintains both a large FoV and high-resolution detail is complicated by the nonrigid motion of human bodies and the subpar pose estimation performance on low-resolution videos. Furthermore, the heterogeneous settings of multi-camera systems, characterized by disparities in FoV, resolution, and viewpoints, make the fusion of human body images a formidable task. Registration and fusion of vastly disparate images are inherently challenging, and this difficulty is exacerbated when dealing with dynamic scenes like moving pedestrians. In such cases, conventional algorithms often yield blurry outputs due to the lack of motion compensation.

A recent study attempted to address this issue by proposing a motion analysis algorithm [2]. By first extracting 3D human poses from both high- and low-resolution videos, high-resolution human motions were then transferred to enhance the inferior pose estimation on low-resolution videos. Despite the success of this approach in transferring human details across hybrid cameras, it lacks a comprehensive mathematical formulation or theoretical analysis. Conversely, contemporary deep learning approaches to human motion transfer often employ neural networks. While these methods demand less in terms of mathematical modeling, they require a substantial

amount of training data and can be time-consuming.

In response to these challenges, we propose a motion transfer algorithm designed specifically for multi-camera systems. Our method presents a comprehensive and explicit mathematical formulation, does not necessitate training data, and offers efficient and interpretable results. Furthermore, we provide a guarantee for motion transfer under the multi-camera setting. Our algorithm extracts the inherent motion patterns from human-centric videos and accomplishes motion transfer tasks using time series analysis. Experimental results on real-world data attest to the effectiveness of our method.

Our main contributions are listed as follows:

- We introduce a time-series-analysis-based motion transfer algorithm for multi-camera systems, complete with a thorough and lucid mathematical formulation.
- Our proposed method is interpretable, eliminates the need for training data, and offers an algorithmic guarantee. In contrast to learning-based techniques, our approach bypasses the need for computationally intensive training procedures.
- Experiments on real-world data and downstream task substantiate the efficacy of our method.

II. RELATED WORK

A. Synthesis of Human Motion

The synthesis of human motion using controlled statistics has been the focus of a substantial body of research [4]–[6]. This work typically utilizes statistical models and learned parameters to generate motions under manual control, thus creating diverse simulations. However, a constraint of these methods is that the diversity of generated motion is tightly coupled with the variability present in the training dataset.

B. Transfer of Human Motion

With the advent of deep learning and neural networks, there has been a shift towards utilizing human pose for predicting

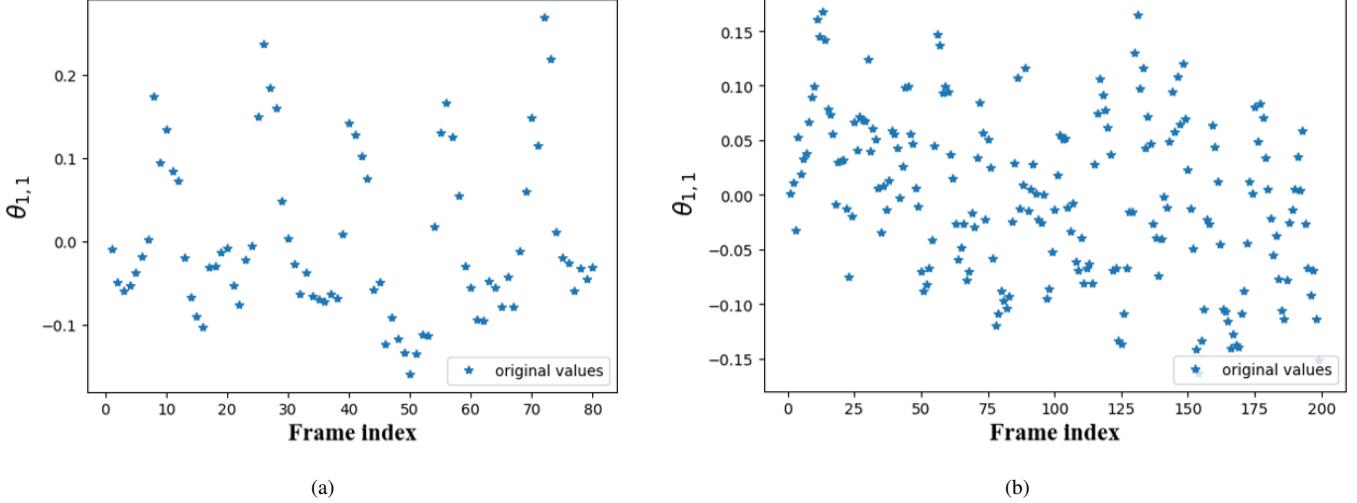


Fig. 3: HR and LR motion data. (a) and (b) represent the $\theta_{1,1}$ value of the HR and LR motion data, respectively.

future frames and synthesizing human videos. For instance, Villegas *et al.* pioneered this trend through their work [7], [8]. Following this, Ma *et al.* synthesized human videos by leveraging a reference image and a target pose [9], [10].

Subsequent studies have advanced this field further. Siarohin *et al.* proposed a deformable network architecture [11]. Chan *et al.* developed a method to extract poses from the source subject and generate the target subject using the learned pose-to-appearance mapping [12]. Liu *et al.* introduced the Liquid Warping GAN to preserve source information [13]. However, all these approaches remain heavily reliant on the training dataset, and their performance significantly deteriorates when applied to real-world data.

C. Time Series Analysis

Time series analysis is a broadly applicable tool, spanning numerous fields [14]–[16]. Geophysics employs it for marine seismic and multi-component streamer data analysis [17], while remote sensing utilizes it for processing Landsat imagery [18]. The field of astronomy leverages it to analyze telescope observations [19], and hydrology uses it for examining streamflow and climate data [20]. Financial industries apply it to analyze stock exchange and price series in daily stock markets [21], and the medical field employs it for processing electroencephalogram and electrocardiogram signals [22].

Despite its widespread use, there has been a conspicuous absence of work that applies time series analysis to the problem of motion transfer between multiple cameras. To bridge this gap, we introduce a motion transfer algorithm for multi-camera systems using time series analysis, with a particular emphasis on seasonality analysis [23], [24].

D. Multi-Camera Systems

The advent of hybrid camera systems was a response to the need for a balance between cameras with different configurations. Among these, dual-camera systems have gained

significant popularity due to their cost-efficiency. Building on this platform, numerous works have pursued image fusion and synthesis to enhance image quality. For instance, Li *et al.* constructed a dual-camera system and proposed a human-centric video super-resolution method that transfers human details from the local-view camera to the global-view camera [2].

A host of other works have built on the dual-camera system. MANet proposed a video denoising approach [25], while Jung *et al.* [26] and Jang *et al.* [27] suggested color transfer techniques. Furthermore, Crossnet [28], Crossnet++ [29], and EFENet [3] proposed methods to perform super-resolution. Yet, despite these advances, the challenge of effectuating efficient, effective, and explainable motion transfer remains unsolved in these systems. Our work aims to address this gap by proposing a new motion transfer algorithm for multi-camera systems.

III. METHODOLOGY

In the context of our multi-camera setting, we propose an algorithm that refines the low-resolution (LR) pose sequence, denoted as θ^L , using time series analysis, with a particular emphasis on seasonality analysis. After obtaining the 3D human poses from high-resolution (HR) and LR videos, the basis of our approach stems from the observation that many human actions, such as walking, running, and physical exercises, exhibit repetitive patterns. Hence, certain human motion data in HR and LR videos can be considered to have seasonal properties.

Our motion transfer algorithm operates through five key steps: (1) identification of seasonality; (2) construction of an additive time series model; (3) detection of periodic points; (4) extraction of the additive factor; and finally, (5) transfer of the motion pattern. For the sake of illustration, we use $\theta_{1,1}$ as an example, which represents the first axis-angle of the first joint in the SMPL model [30] (referring to the ankle joints of the

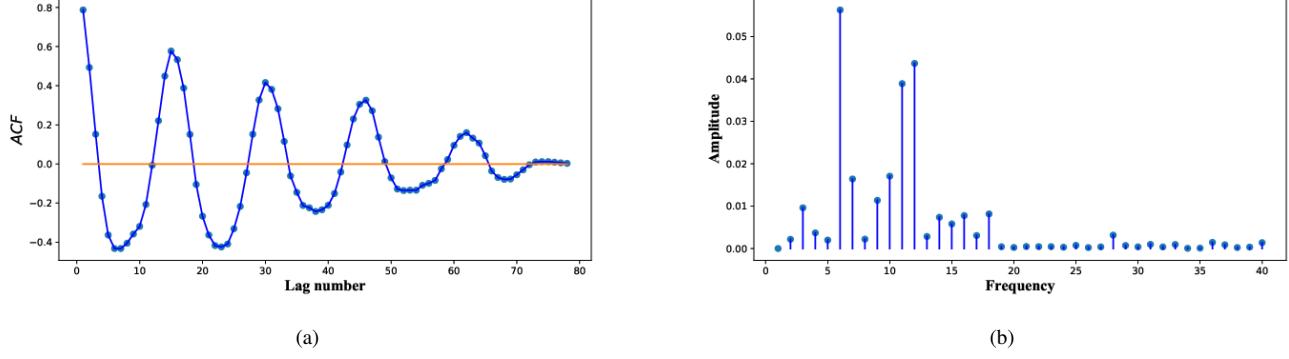


Fig. 4: (a) Auto-correlation result of HR motion data. It performs periodic variation with the maximal value of a function during predefined intervals and gradually decreases to zero. (b) In the Fourier series of the HR motion data, we can see that $f = 5$ has the strongest response, so its reference period is: $80/5 = 16$.

human body). To facilitate subsequent analysis, we normalize the data to the range $[0, 1]$.

Figure 3(a) depicts an example of normalized $\theta_{1,1}$ values in a walking sequence with 80 HR frames, while Figure 3(b) displays the corresponding values for 200 LR frames.

A. Identification of Seasonality

The first step of our process involves the use of the auto-correlation function (ACF) [31] to identify the cyclical nature of HR motion data. As demonstrated in Figure 4(a), the ACF curve exhibits periodic variations, reaching peak values at predefined intervals and eventually decaying to zero.

While human motion tends to be repetitive, the length of a cycle, or period, is not fixed but usually hovers around a constant value. We refer to this value as the “reference period,” denoted as l . It can be estimated using Fourier analysis based on the following relationship:

$$l = n/f, \quad (1)$$

where n denotes the total number of frames in the motion sequence, and f refers to the frequency exhibiting the strongest response. Figure 4(b) presents the Fourier series of the HR motion data. In this instance, we observe that the frequency $f = 5$ shows the strongest response, yielding a reference period of $80/5 = 16$ frames.

B. Construction of Additive Time Series Model

Moving forward, we utilize an additive time-series model [32] to deconstruct the low-resolution pose values into their constituting components. This decomposition is formally expressed as:

$$\mathbf{Y} = \mathbf{S} + \mathbf{T} + \mathbf{E}, \quad (2)$$

where \mathbf{Y} stands for the original pose values, \mathbf{S} symbolizes the short-term variations, \mathbf{T} denotes the long-term trend, and \mathbf{E} represents the noise component. We estimate the long-term trend of the LR sequence, \mathbf{T}^L , by employing polynomial

fitting of the pose values using the least squares method. This fitting can be represented as:

$$\mathbf{T}^L = c_0 + c_1 \theta_{1,1}^L + c_2 (\theta_{1,1}^L)^2 + \cdots + c_{n-1} (\theta_{1,1}^L)^{n-1} + c_n (\theta_{1,1}^L)^n, \quad (3)$$

where c_0, \dots, c_n are constant parameters and $\theta_{1,1}^L$ denotes the $\theta_{1,1}$ value of the LR data. The order n of this fitting function is chosen to be the highest order that ensures $1^{-10} < |c_n| < f$, with c_n representing the coefficients of the highest-order term, and f the number of periods in a pose value sequence.

C. Location of Periodic Points

In the next phase, we initially apply a moving average to smooth the pose values to minimize the interference of noise. The periods can subsequently be identified by locating the crossover points of the original and polynomial fitted value curves, illustrated as red circles in Figure 5(a).

To exclude any unreasonable period indices, we retain only those indices with the period lying within the range $[0.8l, 1.2l]$. Following this, we compute the additive factors [33], denoted as \mathbf{A} , by averaging all the periods. These are then added back to the long-term trend \mathbf{T} to yield the final refined LR poses, as demonstrated in Figure 5(b) and Figure 5(c).

We determine the frame indices corresponding to the crossover points of the smoothed pose curve and the long-term trend curve. Since we are working with discrete series rather than continuous ones, we identify the positive and negative sign changes of the difference between the two series. However, due to the potential imperfections in polynomial fitting, the estimation of \mathbf{T} might exhibit minor fluctuations to account for the extreme values in the pose sequence. To rectify this, we employ the reference period l to exclude any unreasonable period indices. Specifically, for each periodic index p , we validate its authenticity by verifying the presence of a neighboring periodic index p' that adheres to the following

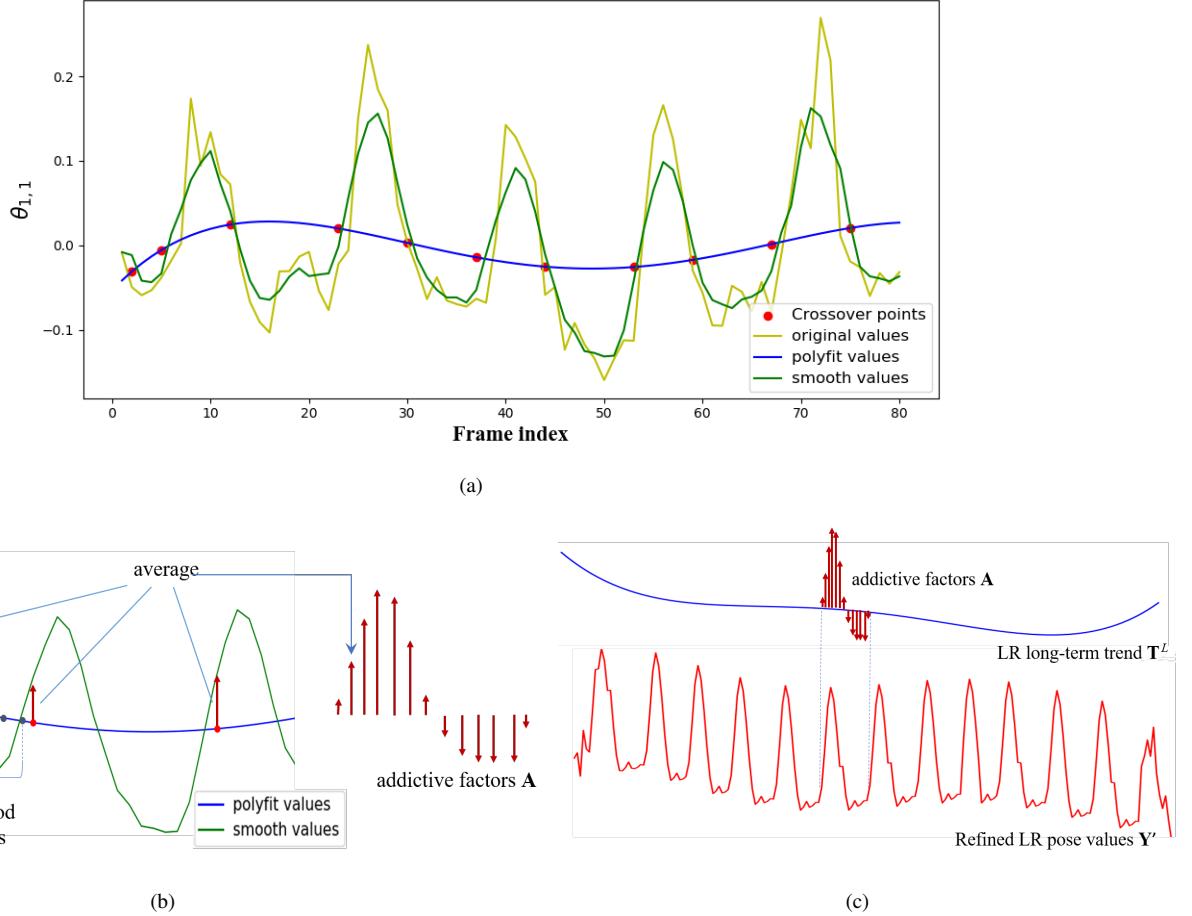


Fig. 5: (a) Crossover points location. The periods could be located by finding the crossover points of the blue and dark green curves, shown as red circles. (b) To refine LR pose values with HR ones, we estimate the additive factor \mathbf{A} by averaging all the periods. (c) To refine LR pose values with HR ones, we add back the additive factor to the long-term trend \mathbf{T}^L to generate the final refined LR poses.

conditions:

$$\begin{aligned} p' &= \arg \min_{p'} |p' - p|, \\ \text{s.t. } &||p' - p| - l| < (1 - \alpha) \times l, \end{aligned} \quad (4)$$

Here, we set the confidence level, α , to 80%, with l representing the reference period. This process ensures that each periodic point p has a neighboring periodic point p' , such that their distance falls within the confidence interval.

D. Extraction of Additive Factor

Upon identifying the periods in the time series, we calculate the minimum period number l_{min} of both HR and LR sequences, and divide each period into l_{min} intervals as uniformly as possible. Then, to extract the additive factor crucial for motion pattern transfer, we remove the long-term trend component from the HR pose values by applying Equation 2:

$$\mathbf{A} = \mathbf{Y}^H - \mathbf{T}^H = \mathbf{S}^H + \mathbf{E}^H, \quad (5)$$

Here, \mathbf{A} represents the additive factor—the discrepancy between the original pose \mathbf{Y}^H and the long-term trend \mathbf{T}^H . It is comprised of the short-term variation \mathbf{S}^H and the noise component \mathbf{E}^H .

Equation 5 allows us to extract the periodically repetitive pattern \mathbf{A} necessary for motion transfer, as depicted in Figure 5(b). As the additive factor \mathbf{A} varies in each period, we utilize the mean value across all periods to execute the motion transfer in a comprehensive manner. Consequently, we infer the mean additive factor $\bar{\mathbf{A}} = (\bar{a}_1, \dots, \bar{a}_i, \dots, \bar{a}_l)$ as:

$$\bar{a}_j = \frac{\sum_{i=1}^n \mathbb{1}\{\phi(i) = j\} a_i}{(f \times m)}, \quad a_i \in \mathbf{A}, \quad (6)$$

In the above equation, a_i refers to the i_{th} value of the additive factor \mathbf{A} , specific to the i_{th} interval in a period; f represents the number of periods; m is the number of values in the i_{th} interval; ϕ is a correspondence function that identifies the frame index i of the time series, within the j_{th} interval of a period:

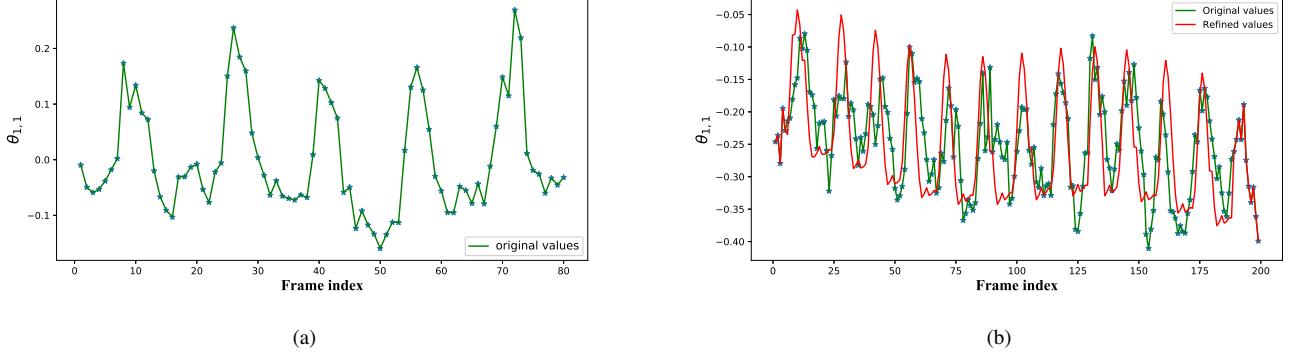


Fig. 6: HR and LR motion data. (a) and (b) represent the $\theta_{1,1}$ value of the HR and LR motion data, respectively.

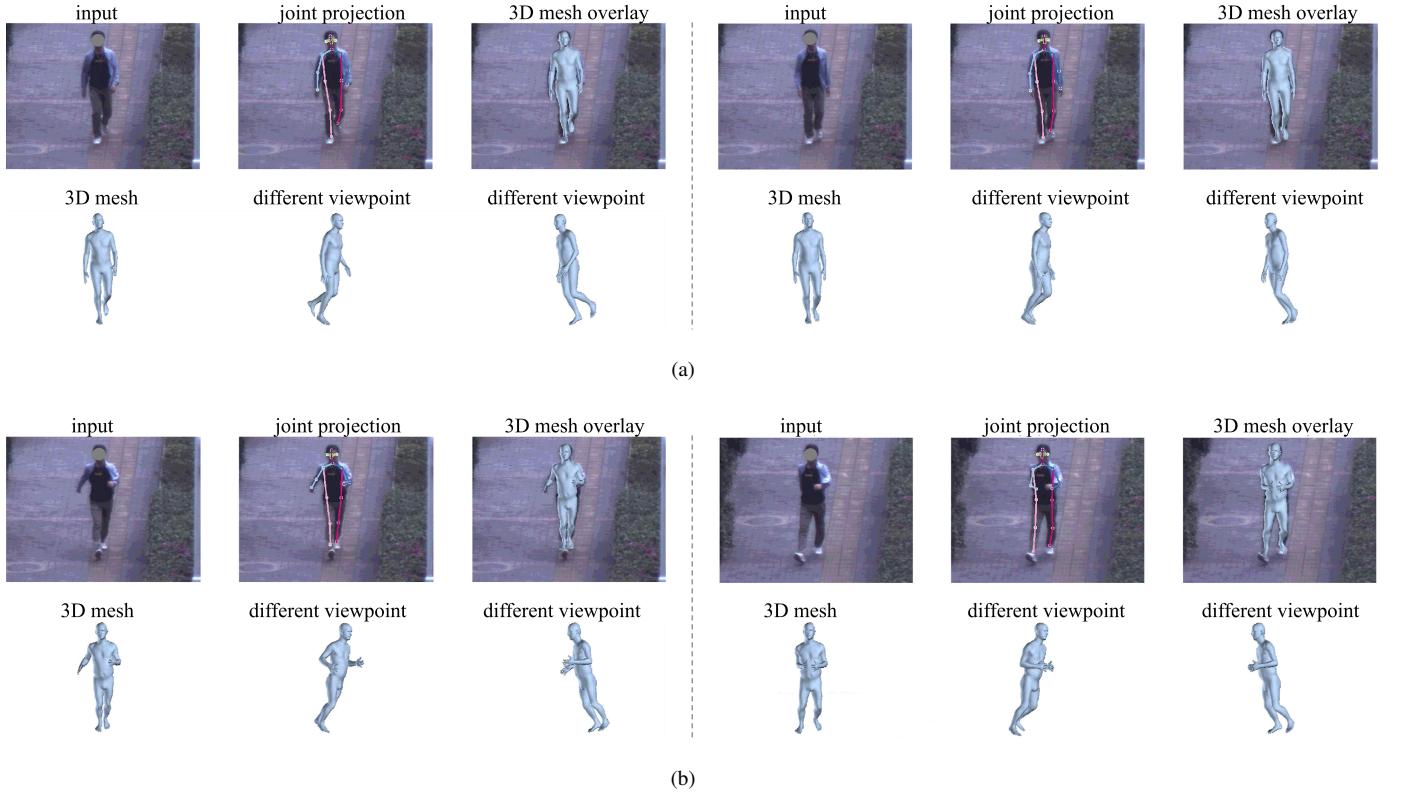


Fig. 7: The SMPL model [34] is adopted to represent 3D body and the HMR [35] is performed to reconstruct 3D human meshes from 2D video frames. (a) and (b) show successful cases and failure cases, respectively. Since we collect and use real-life data in this figure, the human faces are masked out to protect personal privacy.

$$\phi(i) = j. \quad (7)$$

E. Pattern Transfer

In the final phase of our process, we introduce the additive factors, denoted as $\bar{\mathbf{A}}$, to the LR long-term trend, T^L . This integration is implemented with the objective of aligning it with the regular patterns extracted from the high-resolution motion sequence. This step in the process is illustrated in

Figure 5(c). The optimized low-resolution pose values, represented as $\mathbf{Y}' = (y'_1, \dots, y'_i, \dots, y'_n)$, are computed using the following formula:

$$y'_i = t_i + \bar{a}_{\phi(i)}, \quad (8)$$

In this equation, y'_i symbolizes the i_{th} value of the enhanced LR pose, and t_i denotes the i_{th} value of the LR long-term trend T^L . By conducting pattern transfer in this manner, we establish a refined representation of the original low-resolution

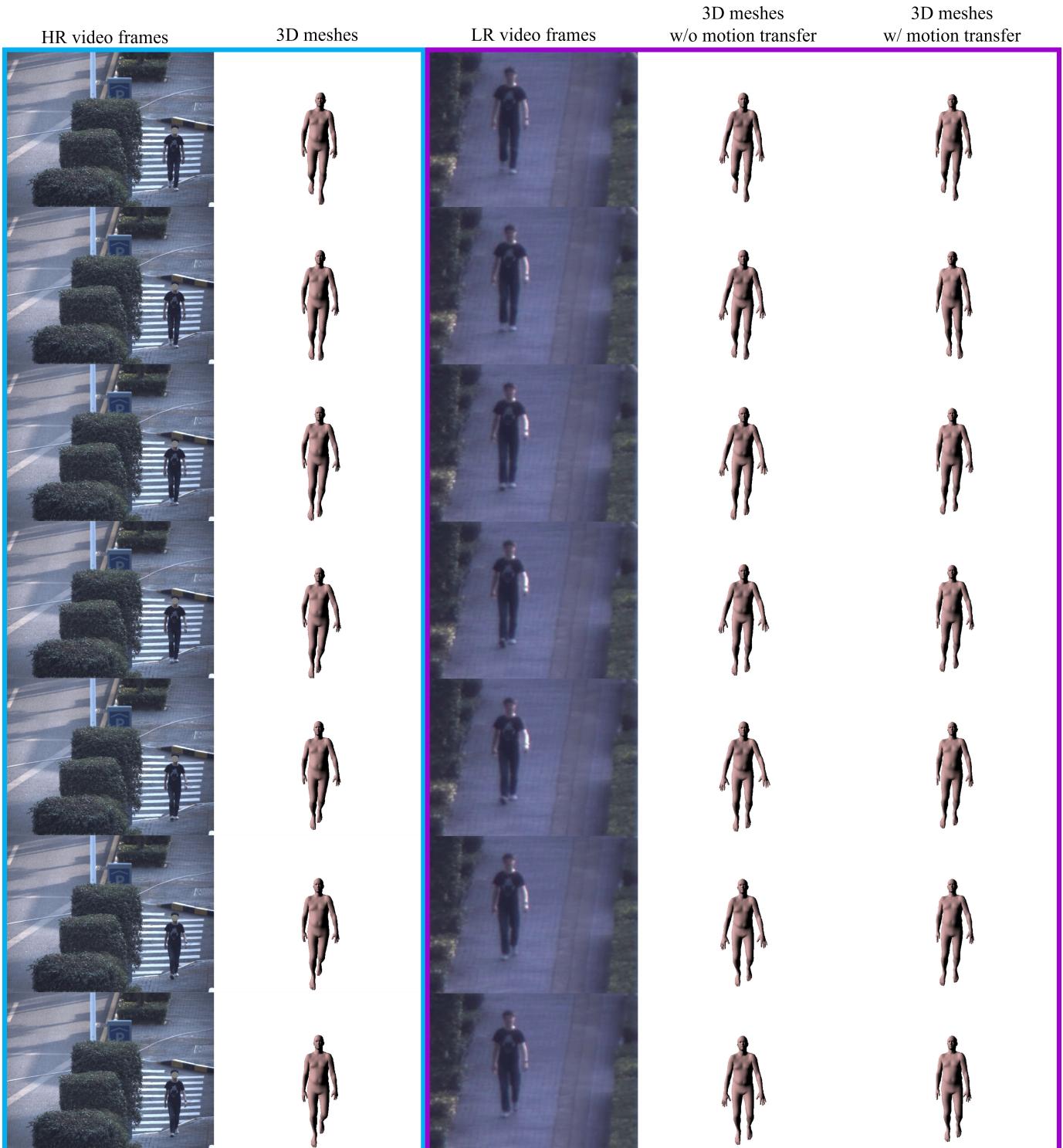


Fig. 8: Performance on 3D human mesh reconstruction using our motion transfer. We observe the weird opening arms due to the low quality of the LR video. In contrast, the results with motion transfer substantially correct the human poses. Since we collect and use real-life data in this figure, the human faces are masked out to protect personal privacy.



Fig. 9: Performance on human detail synthesis across heterogeneous cameras. Top: the synthesized results without our motion transfer method. Bottom: the synthesized results with our method. Without motion transfer, there are intensive jitters on the head and legs. Since we collect and use real-life data in this figure, the human faces are masked out to protect personal privacy.

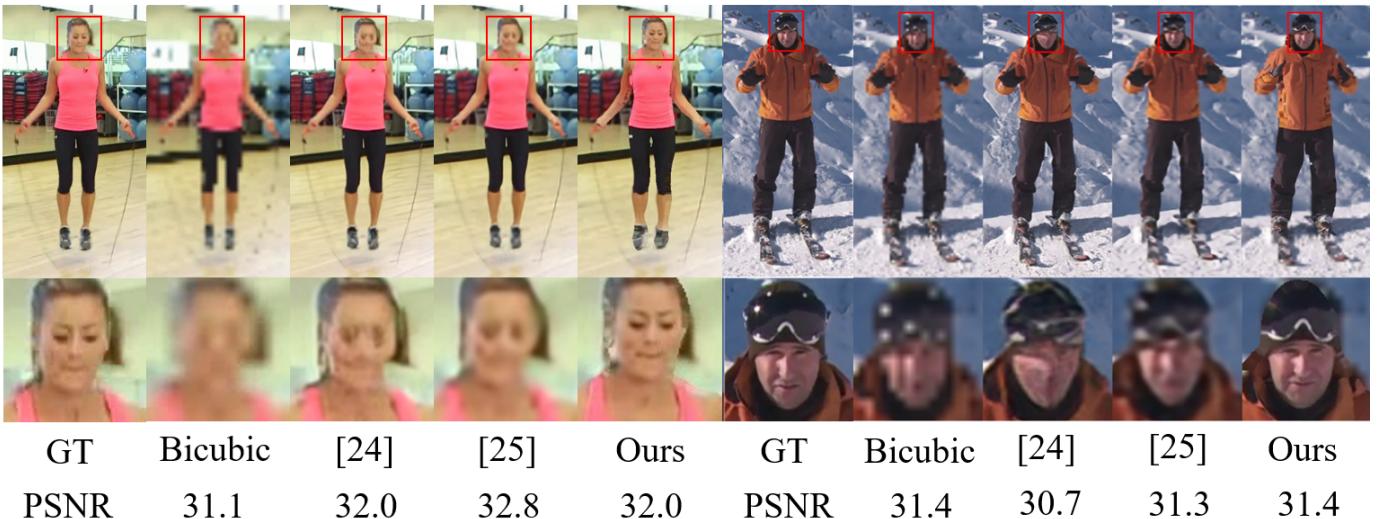


Fig. 10: Results on $\times 8$ RefSR [2] in the synthesised MPII dataset [36]. We benchmark our results against established methods, including Bicubic interpolation and the algorithms proposed by Wang et al. [37], and Tao et al. [38]. Although our method manifests a marginally lower peak signal-to-noise ratio (PSNR), it consistently delivers superior visual quality of the human body in comparison with the other tested methodologies. This underscores the strengths of our approach in maintaining high visual fidelity despite lower quantitative metrics, ultimately emphasizing its potential effectiveness for applications in the domain of computer vision that demand high-quality rendering of human figures.

poses, thereby offering a higher quality approximation of the initial high-resolution sequence.

IV. EXPERIMENT

A. Qualitative Evaluations

In order to validate the effectiveness of our proposed algorithm, we apply it to real-world data procured from a dual-camera system [2]. By using a limited number of high-resolution video frames in conjunction with a complete series of low-resolution video frames, we initially carry out pose estimation using the OpenPose method [39]. This enables us to obtain a sequence of pose values.

As depicted by the green curves in Figure 6, the high-resolution video's estimated pose sequence is of superior

quality. However, the pose sequence extracted from the low-resolution video is noticeably degraded by severe noise. Our motion transfer approach successfully mitigates this issue by refining the low-resolution pose values, as demonstrated in Figure 6(b). Our results effectively preserve the long-term correlation in the low-resolution pose sequence while successfully enhancing the quality of the low-resolution data through the transfer of motion patterns from the high-resolution sequence.

Following the procedures established in [2], we employ the SMPL model [34] to represent 3D bodies and utilize the HMR methodology [35] to reconstruct 3D human meshes from 2D video frames. While we can achieve satisfying results in straightforward cases (for example, where leg movement is minimal and arms are close to knees), as shown in Figure 7(a),

it is clear that HMR struggles to accurately estimate 3D meshes in instances where the subject is walking at a brisk pace and swinging their arms vigorously, as depicted in Figure 7(b).

To enhance the quality of 3D human mesh reconstruction and to further underscore the efficacy of our motion transfer algorithm, we apply our method to high-resolution and low-resolution real-world videos of a pedestrian walking on a street, captured by our dual-camera system. As evidenced in the right of Figure 8, the 3D human meshes reconstructed without our method are inaccurate and implausible, with obvious anomalies such as unnaturally extended arms resulting from the inferior quality of the low-resolution video. However, the use of our motion transfer method results in the accurate correction of these human poses.

Taking our evaluation of the motion transfer method a step further, we integrate it into a downstream vision task. The goal is to embed high-resolution human details into low-resolution videos. As depicted at the top of Figure 9, without our algorithm, the direct synthesis of human details appears inaccurate. For example, obvious jittering can be observed due to the blurred frames, especially in the head and leg regions, with the feet assuming extremely unnatural poses. Conversely, as demonstrated in the bottom row of Figure 9, the application of motion transfer significantly reduces motion jittering, highlighting the utility of our approach.

B. Quantitative Evaluations

Given that our motion transfer approach is applied to the coordinate points of human body joints derived from real-world scenarios, there are no ground truths available, making it challenging to quantify the performance of the motion transfer directly. To address this, we resort to a classic downstream task — video super-resolution through texture transfer — to quantitatively evaluate the proposed motion transfer method.

Specifically, we establish synthetic data sets derived from the MPII dataset [36], which comprises a considerable volume of image sequences centered around human figures and exhibiting a range of human poses. Most MPII sequences consist of about 40 frames. We designate a single frame as the high-resolution frame, while the remaining frames are downsampled by a factor of 8 and treated as low-resolution frames. Following this, we recover the high-resolution details for the downsampled video with the aid of the proposed motion transfer method. Further details regarding the implementation can be found in [2].

As demonstrated in Fig. 10 [2], we conduct a comparative analysis of our method with the methods proposed by Wang et al. [37] and Tao et al. [38]. In the context of synthetic data, as shown in Fig. 10, the approaches in [37] and [38] result in blurry outcomes. In contrast, our method generates visually pleasing results, despite lower PSNR values, especially around the facial region.

V. CONCLUSION

In this study, we present an algorithm for motion transfer among multiple cameras, based on time series analysis. Our

approach has a comprehensive and well-defined mathematical formulation and does not necessitate training data, thereby ensuring efficiency and interpretability. Through rigorous quantitative evaluations on real-world data, we showcase the effectiveness of our method. Furthermore, the versatility of our approach means that it can be extended to accommodate a variety of motion types and is thus not confined to human-centric videos.

REFERENCES

- [1] Xiaoyun Yuan, Lu Fang, Qionghai Dai, David J Brady, and Yebin Liu. Multiscale gigapixel video: A cross resolution image matching and warping approach. In *2017 IEEE International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2017.
- [2] Guanghai Li, Yaping Zhao, Mengqi Ji, Xiaoyun Yuan, and Lu Fang. Zoom in to the details of human-centric videos. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 3089–3093. IEEE, 2020.
- [3] Yaping Zhao, Mengqi Ji, Ruqi Huang, Bin Wang, and Shengjin Wang. Efenet: Reference-based video super-resolution with enhanced flow estimation. In *CAAI International Conference on Artificial Intelligence*, pages 371–383. Springer, 2021.
- [4] Jinxiang Chai and Jessica K Hodgins. Constraint-based motion optimization using a statistical dynamic model. In *ACM SIGGRAPH 2007 papers*, pages 8–es. 2007.
- [5] Jianyu Min and Jinxiang Chai. Motion graphs++ a compact generative model for semantic motion analysis and synthesis. *ACM Transactions on Graphics (TOG)*, 31(6):1–12, 2012.
- [6] Manfred Lau, Ziv Bar-Joseph, and James Kuffner. Modeling spatial and temporal variation in motion data. *ACM Transactions on Graphics (TOG)*, 28(5):1–10, 2009.
- [7] Ruben Villegas, Jimei Yang, Yuliang Zou, Sungryull Sohn, Xunyu Lin, and Honglak Lee. Learning to generate long-term future via hierarchical prediction. In *international conference on machine learning*, pages 3560–3569. PMLR, 2017.
- [8] Ruben Villegas, Jimei Yang, Duygu Ceylan, and Honglak Lee. Neural kinematic networks for unsupervised motion retargetting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8639–8648, 2018.
- [9] Liqian Ma, Xu Jia, Qianru Sun, Bernt Schiele, Tinne Tuytelaars, and Luc Van Gool. Pose guided person image generation. *arXiv preprint arXiv:1705.09368*, 2017.
- [10] Liqian Ma, Qianru Sun, Stamatios Georgoulis, Luc Van Gool, Bernt Schiele, and Mario Fritz. Disentangled person image generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 99–108, 2018.
- [11] Aliaksandr Siarohin, Enver Sangineto, Stéphane Lathuilière, and Nicu Sebe. Deformable gans for pose-based human image generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3408–3416, 2018.
- [12] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A Efros. Everybody dance now. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5933–5942, 2019.
- [13] Wen Liu, Zhixin Piao, Jie Min, Wenhan Luo, Lin Ma, and Shenghua Gao. Liquid warping gan: A unified framework for human motion imitation, appearance transfer and novel view synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5904–5913, 2019.
- [14] W Cleveland. Analysis and forecasting of seasonal time series. 1973.
- [15] George EP Box and Gwilym M Jenkins. Time series analysis. forecasting and control. In *Holden-Day Series in Time Series Analysis, Revised ed., San Francisco: Holden-Day*, 1976, 1976.
- [16] Robert Goodell Brown. *Smoothing, forecasting and prediction of discrete time series*. Courier Corporation, 2004.
- [17] Ebrahim Ghaderpour, Wenyuan Liao, and Michael P Lamoureux. Antileakage least-squares spectral analysis for seismic data regularization and random noise attenuation. *Geophysics*, 83(3):V157–V170, 2018.
- [18] M Razu Ahmed, Quazi K Hassan, Masoud Abdollahi, and Anil Gupta. Introducing a new remote sensing-based model for forecasting forest fire danger conditions at a four-day scale. *Remote Sensing*, 11(18):2101, 2019.

- [19] Ebrahim Ghaderpour and Shahnaz Ghaderpour. Least-squares spectral and wavelet analyses of v455 andromedae time series: The life after the super-outburst. *Publications of the Astronomical Society of the Pacific*, 132(1017):114504, 2020.
- [20] Victor B Veiga, Quazi K Hassan, and Jianxun He. Development of flow forecasting models in the bow river at calgary, alberta, canada. *Water*, 7(1):99–115, 2015.
- [21] Anindya Chakrabarty, Anupam De, Angappa Gunasekaran, and Rameshwar Dubey. Investment horizon heterogeneity and wavelet: Overview and further research directions. *Physica A: Statistical Mechanics and its Applications*, 429:45–61, 2015.
- [22] Abhijit Bhattacharya, Manish Sharma, Ram Bilas Pachori, Pradip Sircar, and U Rajendra Acharya. A novel approach for automated detection of focal eeg signals using empirical wavelet transform. *Neural Computing and Applications*, 29(8):47–57, 2018.
- [23] Svend Hylleberg. *Modelling seasonality*. Oxford University Press, 1992.
- [24] Kenneth F Wallis. Seasonal adjustment and relations between variables. *Journal of the American Statistical Association*, 69(345):18–31, 1974.
- [25] Yaping Zhao, Haitian Zheng, Zhongrui Wang, Jiebo Luo, and Edmund Y Lam. Manet: Improving video denoising with a multi-alignment network. *arXiv preprint arXiv:2202.09704*, 2022.
- [26] Yong Ju Jung. Enhancement of low light level images using color-plus-mono dual camera. *Optics express*, 25(10):12029–12051, 2017.
- [27] Hae Woong Jang and Yong Ju Jung. Deep color transfer for color-plus-mono dual cameras. *Sensors*, 20(9):2743, 2020.
- [28] Haitian Zheng, Mengqi Ji, Haoqian Wang, Yebin Liu, and Lu Fang. Crossnet: An end-to-end reference-based super resolution network using cross-scale warping. In *Proceedings of the European conference on computer vision (ECCV)*, pages 88–104, 2018.
- [29] Yang Tan, Haitian Zheng, Yinheng Zhu, Xiaoyun Yuan, Xing Lin, David Brady, and Lu Fang. Crossnet++: Cross-scale large-parallax warping for reference-based super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12):4291–4305, 2020.
- [30] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. *ACM transactions on graphics (TOG)*, 34(6):1–16, 2015.
- [31] Charles I Plosser. Short-term forecasting and seasonal adjustment. *Journal of the American Statistical Association*, 74(365):15–24, 1979.
- [32] Agustín Maravall and David A Pierce. A prototypical seasonal adjustment model. *Journal of Time Series Analysis*, 8(2):177–193, 1987.
- [33] Mohammed Yar and Chris Chatfield. Prediction intervals for the holt-winters forecasting procedure. *International Journal of Forecasting*, 6(1):127–137, 1990.
- [34] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. *ACM transactions on graphics (TOG)*, 34(6):248, 2015.
- [35] Angjoo Kanazawa, Michael J Black, David W Jacobs, and Jitendra Malik. End-to-end recovery of human shape and pose. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7122–7131, 2018.
- [36] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3686–3693, 2014.
- [37] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. EsrGAN: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 0–0, 2018.
- [38] Xin Tao, Hongyun Gao, Renjie Liao, Jue Wang, and Jiaya Jia. Detail-revealing deep video super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4472–4480, 2017.
- [39] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Openpose: realtime multi-person 2d pose estimation using part affinity fields. *IEEE transactions on pattern analysis and machine intelligence*, 43(1):172–186, 2019.