Australasian Transport Research Forum 2023 Proceedings
29 November – 1 December, Perth, Australia
Publication website: http://www.atrf.info

Improving efficiency of origin-destination estimation problem through keyframe interpolation

Raghav Malhotra[1], Chintan Sanjeev Advani[2], Paul Corry[3], Ashish Bhaskar[4]
[1]raghav.malhotra@hdr.qut.edu.au
[2]chintan.advani@hdr.qut.edu.au
[3]p.corry@qut.edu.au
[4]ashish.bhaskar@qut.edu.au
Email for correspondence: ashish.bhaskar@qut.edu.au

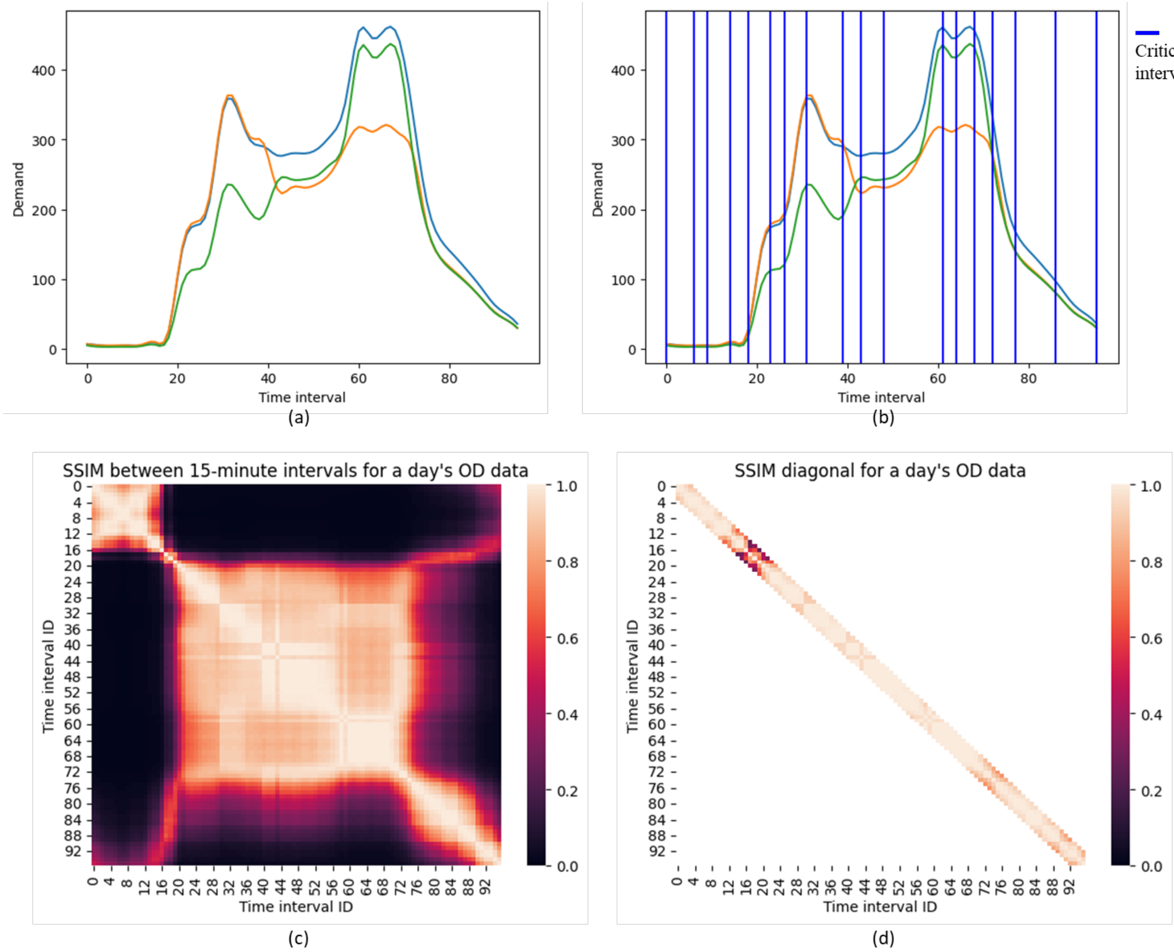
1. Introduction

The Dynamic Origin-Destination estimation (DODE) problem is a well-documented problem in academic literature. The relationship between traffic demand (Origin-Destination matrices) and traffic states (link counts) is quantified using a system of linear equations. Here, the mapping between the OD and link counts is empirically expressed using the assignment matrix. However, the imbalance in the number of unknowns (OD pairs) and the number of measurements (link counts) lead to the problem being under-determined as the number of OD pairs are generally much larger than the number of links. Therefore, the problem is generally solved using optimisation techniques. In addition to being under-determined, the OD estimation problem also has high dimensionality, especially in large urban networks wherein the number of OD pairs can span from several tens of thousands to millions. This makes the problem of OD estimation, a non-trivial one.

The above-mentioned problems in DODE have been tackled in the past by either reducing the number of unknowns (Djukic et al., 2012; Krishnakumari et al., 2019; Lorenzo & Matteo, 2013) or incorporating additional measurements in the form of other data sources such as speed, density, and occupancy (Balakrishna et al., 2007; Qurashi et al., 2019; Tympakianaki et al., 2015). In the solution frameworks proposed in the above-mentioned literature, researchers have approached the problem of high temporal complexity of DODE by either reducing the problem size in its space domain, or focusing on a quasi-dynamic methodology which assumes the within-day dynamics of the OD to be constant.
In this paper, a different approach of reducing the temporal dimensionality is proposed for the adjustment of current time-dependent OD matrices. We propose limiting the number of intervals to be estimated under the assumption that the OD demand as a time series, progresses gradually through the day, rather than possessing abrupt changes. This means that the identification and estimation of a subset of the time intervals (key intervals) can facilitate the overall OD estimation process by interpolating the non key intervals. To demonstrate this, consider Figure 1 (c) and (d) which shows the heatmap of the Structural Similarity Index Measure (SSIM) between all intervals. In Figure 1 (c), we note that the intervals are structurally similar within parts of the day but much different in comparison to distant intervals such as, morning peak intervals have no similarity with the inter-peak or afternoon peak OD matrices. On the other hand, Figure 1 (d) can also show that the progression of the structure of OD is gradual.
Figure 1 (a) shows three time series data for motorway OD pairs. A sample of intervals are labelled using the vertical lines in Figure 1 (b). This gives a premise to the proposed methodology for time-efficient DODE where only the key intervals for each OD pair are estimated and the gaps can be interpolated which reduces the computational complexity of the estimation process many-fold while having minimal effect on the accuracy.
It is to be noted that the methodology is data source independent, i.e., any source of OD and traffic state data can be used as long as the solution framework allows it. For example, a generalised least square formulation of the OD optimization problem, solved using the gradient descent method requires an

empirical relationship between the independent (OD matrices) and the dependent (traffic states such as traffic flow) datasets.

Figure 1: (a) OD demand for sample motorway OD pairs, (b) OD demand with its critical points, (c) SSIM heatmap for day's OD data, and (d) Diagonal representation of OD demand



2. Methodology

The motivation behind keyframe interpolation technique for OD adjustment comes from the field of video summarization, in which the structure of frames which make up a video are analysed to extract a subset (known as key-frames) of frames which can in-turn be used to summarize a given video. A video is made up of a series of frames and a small collection of similarly structured frames is known as a shot. The first process of keyframe summarization is to identify the boundary frames of these shots, also known as the process of shot boundary detection (SBD). Shot boundaries in the video context can be either abrupt or gradual. This method is then followed by key-frame extraction which involves selecting one or more than one frames from each video shot which can be used a representative frame.

In the OD estimation context, the key idea remains the same where we identify the key intervals in a series of temporal OD matrices. This is followed by a standard OD adjustment process of these key intervals and the missing ODs in between are filled using an interpolation technique.

2.1. Shot boundary detection (SBD)

Shot Boundary Detection (SBD) is the process of identifying critical intervals to be adjusted and interpolated. SBD is a preprocessing task which utilises the prior OD matrices to identify which intervals in the time horizon define the overall trends in the demand progression through the day. There are several techniques with which this can be achieved which are detailed below.

2.1.1. Rule based detection.

In rule-based classifiers, the classification of a boundary or non-boundary frames is based on a threshold. This is further classified into two categories as **Simultaneous interval** and **linear discontinuity search**. In the case of simultaneous interval, we check for a boundary by analysing the similarity between consecutive intervals. The classification is given by:

$$\omega_t = \begin{cases} 0, & if\ s_t > T \\ 1, & otherwise \end{cases} \tag{1}$$

Where T is a predefined threshold, ω_t is the classification parameter for boundary intervals where $\omega_t = 0$ means it is a boundary and 1 meaning it isn't. s_t is defined as the similarity index between time t and $t+1$.

Linear discontinuity search builds upon the simultaneous interval by not only check for the similarity between t and $t+1$, but also between t and $t+2, t+3, t+4$ and so on until the condition is satisfied by (1).

2.1.2. Statistical machine learning

In the past few years, there have been several attempts at picturing the problem of SBD as a pattern recognition problem. Yuan et al. (2007) categorizes the statistical machine learning solution frameworks for SBD into two sub-categories, i.e., Generative classifiers and discriminative classifiers. Generative classifiers can describe the statistics behind their classifications while discriminative classifiers do not. Generative classifiers not only allow but are highly dependent on prior information provided to the model, as shown by (Vasconcelos & Lippman, 2000), (Hanjalic, 2002) and (Janvier et al., 2003). In the video summarization process, various generative and discriminative classifiers have been applied in the literature such as:

1. Generative classifiers:
   - Generative adversarial networks
   - Variational auto-encoders
2. Discriminative classifiers:
   - K-means clustering (Naphade et al., 1998)
   - K-nearest neighbours (Cooper, 2004)
   - Support vector machines (Yuan, Li, et al., 2005), (Ngo, 2003), (Chua et al., 2003)

2.1.3. Graph based

Yuan, Zhang, et al. (2005) proposed a novel shot boundary detection method taking the inspiration from photo segmentation using a graph partition model for temporal data segmentation. Segmentation by graph cuts requires one to define the temporal data as a weighted graph structure *G(V, E)* with a set of nodes *V* and a set of edges *E* where the weights can be defined using various metrics such as correlation, structural similarity etc. This is followed by calculating a *score(t)* for each $t \in \tau$ which defines the cost of making a *cut* at the time interval t. The time interval t for which $score(t)$ is minimum is selected which leads to classification of $\{1,2,3 \ldots t\}$ and $\{t+1, t+2, t+3 \ldots, T\}$ as the clusters.

In the OD adjustment context, OD is also represented as a temporal data structure and the weights $w(i,j)$ of *G(V, E)* are defined as follows:

$$w(i,j) = SSIM(i,j) \tag{2}$$

Where $SSIM(i,j)$ is the structural similarity index between the intervals i and j.

This process of clustering follows the concept of maximising intra-cluster dissimilarity and inter-cluster similarity. Yuan, Zhang, et al. (2005) defines $cut(A,B)$, $assoc(A)$ and $score(t)$ as follows:

$$cut(A,\ B) = \sum_{i \in A,\ j \in B} w_{ij} \tag{3}$$

$$assoc(A) = \sum_{i,\, j \in A} w_{ij} \tag{4}$$

$$score(t) = \frac{cut(A,\, B)}{assoc(A)} + \frac{cut(A,\, B)}{assoc(B)} \tag{5}$$

$cut(A, B)$ is defined as the cost of dividing the time series into clusters A and B,
$assoc(A)$ is defined as the association of all links within cluster A , and
$score(t)$ is defined as the overall score of dividing the time series into clusters A and B.
This process is an iterative process as after each cut is made, we can further segment it until a desirable cluster size is reached.

2.2. Keyframe extraction

One we have clustered our entire temporal structure, each of these clusters leads to one or more than one key-interval (frame) which are used as the representative for the cluster.
This task can be achieved by selecting any of the following intervals from each cluster:

- First interval
- Last interval
- Central interval
- First and last interval

For OD adjustment, this research uses first interval from each cluster as the representative interval.

2.3. Interpolation

Keyframe extraction process is followed by the OD estimation process for those key intervals independently. Bi-level approach using generalised least square formulation is used to define the optimization problem in this research, which is solved using gradient descent approach proposed by (Ros-Roca et al., 2021).
The adjusted key-intervals are not the complete solution to the OD adjustment problem. An interpolation process is used to fill the missing OD matrices. Several interpolation techniques exist in literature and the following are used for testing purposes for keyframe interpolation method.

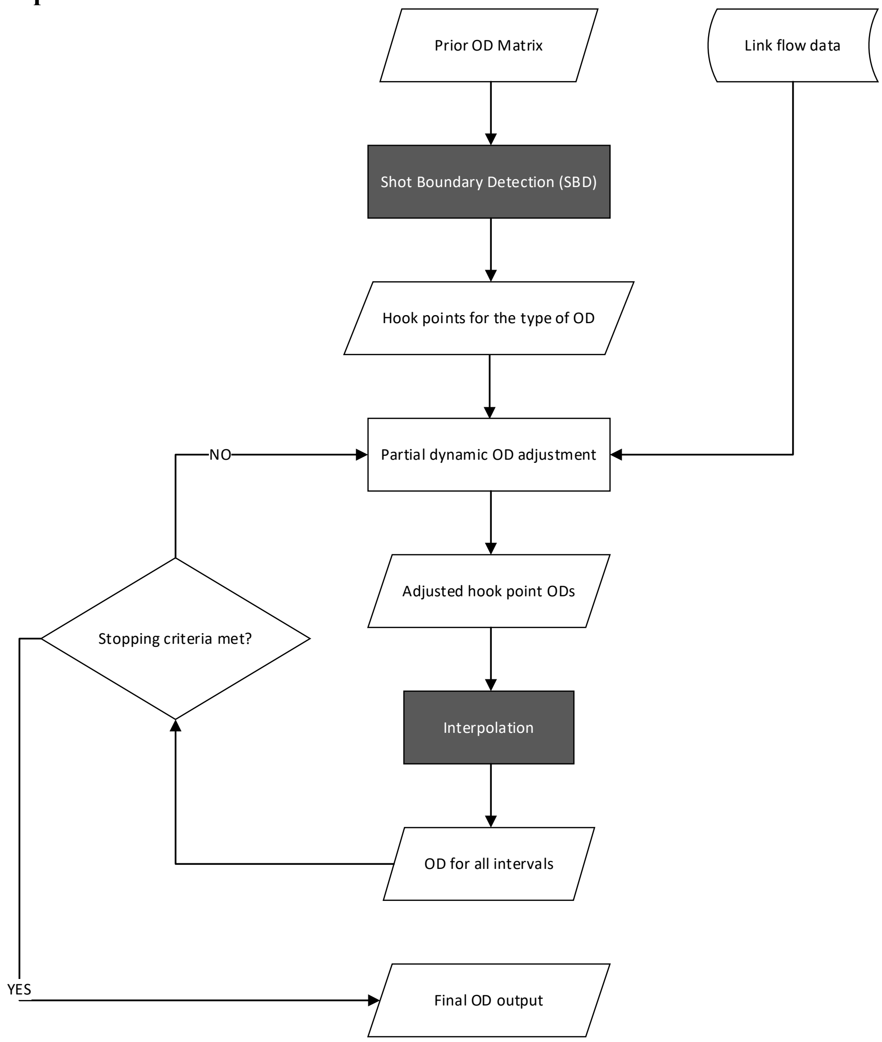
- Linear interpolation
- Polynomial interpolation
- Spline interpolation
- Fourier interpolation

The first three interpolation techniques work under the premise of using the available information to fit a function(s) through the known points. In the case of Fourier interpolation, the known points are used to find the number of wavelets to be used to reconstruct OD time series for an OD pair. Minimising error among the known points gives us the number of wavelets to be used to apply an inverse Fourier transformation. This results in a final OD output which consist of OD matrices for all time intervals. Fourier interpolation has the property of utilising the prior OD information and patterns to fill the missing gap, therefore, a "good" knowledge in the OD flow patterns can lead to better interpolation results.

2.4. Framework

In an OD estimation process using bi-level approach, the lower level (assignment through simulation), takes the most time as compared to other subtasks. Keyframe interpolation makes improving this runtime possible by avoiding running a simulation for all intervals. Depending on the network characteristics, smaller simulations only for the key intervals are run, which leads to faster convergence at the lower level. It can be further enhanced by running these simulations parallelly as they are independent of each other. The framework for KFI (Keyframe Interpolation) method is shown in Figure 2.

Figure 2: KFI with parallel simulations



3. Results and conclusions

The performance of KFI approach for OD estimation is benchmarked on the Logan city area network which consists of 284 centroids (80,656 OD pairs per time interval) and 441 loop detector stations. The new methodology is benchmarked against the state-of-art bi-level approach for OD estimation. The ground truth OD is provided by Aimsun as representative for a working day traffic demand, representing . A uniform noise is then added to this ground truth OD using the methodology specified by (Antoniou et al., 2016) which is used as a prior/base OD. Ground truth network states are generated by simulating the ground truth traffic demand.
For the purpose of this study, above-mentioned methodologies are evaluated on two metrics:
1. Structural similarity index measure (SSIM)
2. GEH statistic [1]
3. Runtime

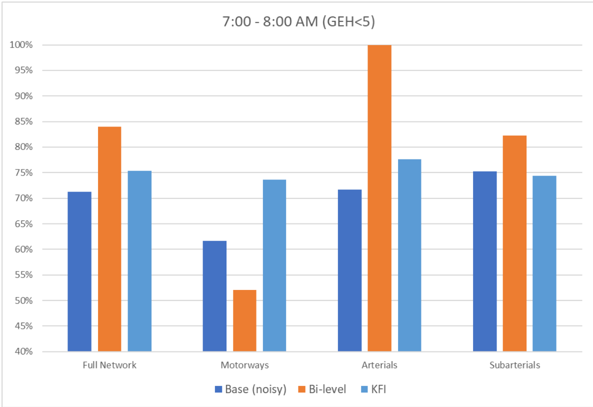
The highlighted intervals in Figure 2 (a) represent the key intervals identified from the analysis period using graph partition model. We see an overall improvement in SSIM metric for both KFI and bi-level adjustment, however, KFI method seems to outperform bi-level adjustment method.

1 Geoffrey E. Havers invented the GEH statistic in 1970, which is named after him.

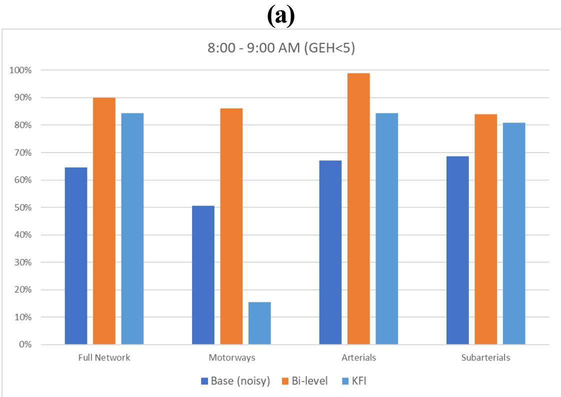
Figure 3: (a)SSIM results and (b)-(d) GEH results for KFI vs Bi-level.

Interval	Base	BiLevel	KFI
1	0.77	0.77	0.98
2	0.78	0.87	0.82
3	0.8	0.83	0.93
4	0.76	0.78	0.9
5	0.68	0.71	0.98
6	0.73	0.77	0.98
7	0.74	0.72	0.75
8	0.78	0.83	0.99
9	0.77	0.87	0.88
10	0.68	0.76	0.99
11	0.79	0.89	0.99
12	0.71	0.79	0.81

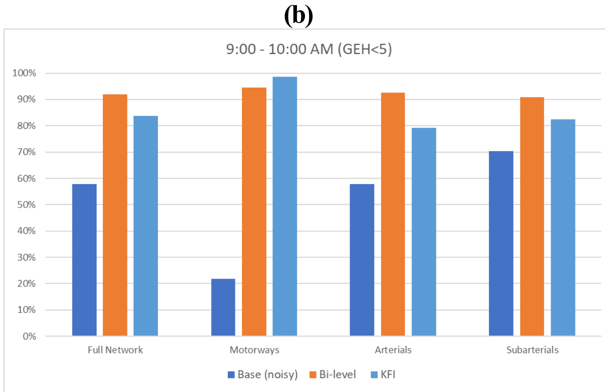
(a)



(b)



(c)

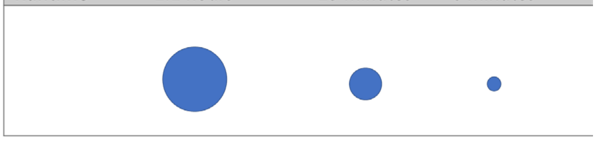


(d)

This is highly likely a result of the Fourier transform being applied on a small deviation of true OD. In principle, these results get affected directly by how good OD prior knowledge is. GEH results have been represented for each hour of the analysis period as well by the road hierarchy in Figure 2 (b) - (d). We see that overall, both KFI and bi-level show similar results while KFI performs much better than bi-level for motorways. Bi-level adjustment shows a much better performance on arterials.

Figure 4: Runtime complexity results

Adjustment technique	Bi-level	KFI with split simulation	KFI with split parallel simulations*
Runtime	2.1 hours	25 minutes	6 minutes



Runtime is an attractive metric to evaluate as OD matrix adjustment is computationally very expensive. Table 10 shows that for the same given problem, KFI can reach convergence in a much faster runtime while maintaining similar accuracy. KFI with parallel simulations runtime is an expected and not tested runtime because of the lack of the feature to run parallel simulations currently.

4. References

Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., Ciuffo, B., Djukic, T., Hoogendoorn, S., & Marzano, V. (2016). Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation. *Transportation Research Part C: Emerging Technologies*, *66*, 79-98.

Balakrishna, R., Ben-Akiva, M., & Koutsopoulos, H. N. (2007). Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation. *Transportation Research Record*, *2003*(1), 50-58.

Chua, T.-S., Feng, H., & Chandrashekhara, A. (2003). An unified framework for shot boundary detection via active learning. 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).

Cooper, M. (2004). Video segmentation combining similarity analysis and classification. Proceedings of the 12th annual ACM International Conference on Multimedia,

Djukic, T., Flötteröd, G., Van Lint, H., & Hoogendoorn, S. (2012). Efficient real time OD matrix estimation based on Principal Component Analysis. 2012 15th International IEEE Conference on Intelligent Transportation Systems,

Hanjalic, A. (2002). Shot-boundary detection: unraveled and resolved? *IEEE transactions on circuits and systems for video technology*, *12*(2), 90-105.

Janvier, B., Bruno, E., Marchand-Maillet, S., & Pun, T. (2003). Information-theoretic framework for the joint temporal partionning and representation of video data. Proceedings of the 3rd International Workshop on Content-Based Multimedia Indexing, CBMI'03,

Krishnakumari, P., van Lint, H., Djukic, T., & Cats, O. (2019). A data driven method for OD matrix estimation. *Transportation Research Procedia*, *38*, 139-159.

Lorenzo, M., & Matteo, M. (2013). OD matrices network estimation from link counts by neural networks. *Journal of Transportation Systems Engineering and Information Technology*, *13*(4), 84-92.

Naphade, M. R., Mehrotra, R., Ferman, A. M., Warnick, J., Huang, T. S., & Tekalp, A. M. (1998). A high-performance shot boundary detection algorithm using multiple cues. Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269),

Ngo, C.-W. (2003). A robust dissolve detector by support vector machine. Proceedings of the eleventh ACM international conference on Multimedia,

Qurashi, M., Ma, T., Chaniotakis, E., & Antoniou, C. (2019). PC–SPSA: employing dimensionality reduction to limit SPSA search noise in DTA model calibration. *IEEE Transactions on Intelligent Transportation Systems*, *21*(4), 1635-1645.

Ros-Roca, X., Montero, L., & Barceló, J. (2021). Investigating the quality of Spiess-like and SPSA approaches for dynamic OD matrix estimation. *Transportmetrica A: Transport Science*, *17*(3), 235-257.

Tympakianaki, A., Koutsopoulos, H. N., & Jenelius, E. (2015). c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation. *Transportation Research Part C: Emerging Technologies*, *55*, 231-245.

Vasconcelos, N., & Lippman, A. (2000). Statistical models of video structure for content analysis and characterization. *IEEE Transactions on Image Processing*, *9*(1), 3-19.

Yuan, J., Li, J., Lin, F., & Zhang, B. (2005). A unified shot boundary detection framework based on graph partition model. Proceedings of the 13th annual ACM international conference on Multimedia,

Yuan, J., Wang, H., Xiao, L., Zheng, W., Li, J., Lin, F., & Zhang, B. (2007). A Formal Study of Shot Boundary Detection. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(2), 168-186. https://doi.org/10.1109/TCSVT.2006.888023

Yuan, J., Zhang, B., & Lin, F. (2005). Graph partition model for robust temporal data segmentation. Advances in Knowledge Discovery and Data Mining: 9th Pacific-Asia Conference, PAKDD 2005, Hanoi, Vietnam, May 18-20, 2005. Proceedings 9