Estimation of Transfer Coefficients and Signals of Sound-to-Light Conversion Device Blinky Under Saturation

Kosuke Nishida*, Natsuki Ueno*, Yuma Kinoshita*†, and Nobutaka Ono*
* Tokyo Metropolitan University, Tokyo, Japan
E-mail: nishida-kosuke@ed.tmu.ac.jp
† Tokai University, Kanagawa, Japan

***Abstract*—In this study, we propose a method for estimating the intensity of the light signal emitted by a sound-to-light conversion device called Blinky from the saturated video signal obtained by a camera. A Blinky is a compact device that converts acoustic information into the light signal, and distributed Blinkies are applicable to various acoustic sensing frameworks without wired or wireless communication by monitoring them with a single video camera. Here, it is desired to obtain the intensity of the output light signal of the Blinky from the observed video signal; however, these are generally different owing to several factors such as light attenuation, background light, and saturation of brightness in the camera. We first calibrate coefficients describing the relationship between these signals using a given reference signal and then estimate the unknown signal of the Blinky using the calibrated coefficients. In each step, we consider the saturation effect in the observed video signal, which is a main contribution of this study. Experimental evaluations were performed under simulated and real environments, and their results indicated that the proposed method achieved higher estimation accuracy than the conventional method without consideration of the saturation effect.**

I. Introduction

Acoustic signal processing using distributed microphones has a wide variety of applications such as source localization [1], [2], audio source separation [3], [4], and acoustic scene classification [5], [6]. Especially, the use of microphones distributed arbitrarily over space without wired or wireless communication is challenging but practically useful in various situations. In this problem, the main difficulty lies in synchronization of signals observed by multiple microphones.

To make synchronization easy, a new sensing framework using sound-to-light conversion devices called *Blinkies* was proposed by Scheibler and Ono [7]. A Blinky is a compact, battery-powered device equipped with a microphone, light-emitting diode (LED), and microcomputer as shown in Fig. 1 (left), where an acoustic signal captured by the microphone is converted to a light signal of the LED via a programmable conversion process. By observing distributed Blinkies with a single video camera, it is easy to synchronize the signals of many Blinkies unconnected with each other.

Blinkies have been applied in various types of acoustic signal processing including audio source separation [8], multiple source localization [9], real-time pitch analysis [10], and acoustic scene analysis [11]. In some of these applications, it is necessary to obtain the intensity of the output light signals of Blinkies (referred to as the Blinky signals, hereafter). However, it is not always an easy task because the brightness of the observed video signals (referred to as the observed signals, hereafter) is different from the intensity of the Blinky signals owing to several factors such as light attenuation, background light, and saturation of brightness in a camera. Therefore, calibration, i.e., quantification of relationship between Blinky signals and observed signals, has to be performed first using given reference Blinky signals, and then the unknown Blinky signals can be restored from the observed signals based on the calibrated relationship. The calibration and signal restoration methods were proposed in [9]; however, saturation of brightness in a camera is yet to be taken into consideration.

In this study, we present new calibration and signal restoration methods considering saturation of the observed signals. The main idea is based on the fact that the Blinky signals are observed in multiple pixels for different brightness values, which is a distinct feature in this situation compared to other general situations of restoration problem of saturated signals [12]–[14]. We integrate information of nonsaturated pixels and formulate least-squares problems for calibration and signal restoration. Experimental evaluations were performed under simulated and real environments, and their results indicated that the proposed method achieved higher estimation accuracy than the conventional method [9] in a condition where the observed signals were saturated.

II. Problem Setting

A. Overview

Suppose multiple Blinkies are distributed and monitored by a single video camera as shown in Fig. 1 (right). As described in Sect. I, the observed signal is different from the Blinky signal, which may cause difficulties in direct application to some types of post-stage processing. The positions of the Blinkies and camera are assumed to be fixed in calibration and signal estimation, but they need not be given. Here, our objective is twofold: to calibrate coefficients describing the relationship between them using a given reference Blinky

978-616-590-477-3 ©2022 APSIPA APSIPA ASC 2022

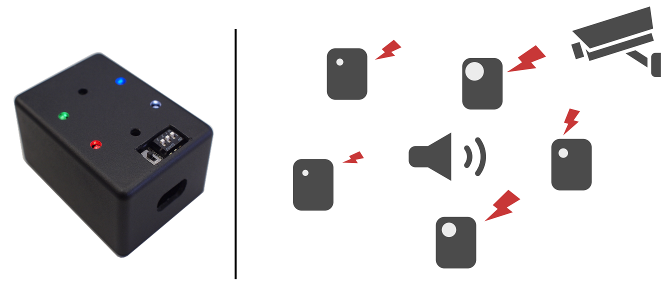


Fig. 1: Sound-to-light conversion device Blinky (left) and acoustic sensing framework using Blinkies and single video camera (right).

signal and to estimate the unknown Blinky signal using the calibrated coefficients. In this study, we focus on a single Blinky; however, the proposed method can be used to multiple Blinkies independently under the assumption that the light signals from different Blinkies do not overlap in the observed pixels.

B. Observation Model

Let s_n and $p_{i,n}$ denote the Blinky signal and the signal observed by the video camera, respectively, where i and n is the pixel and time index, respectively. Although the pixels are arranged in a two-dimensional grid, a single index i is used here because there is no need to distinguish the vertical and horizontal dimensions in the proposed framework. Then, the relationship between s_n and $p_{i,n}$ is given by

$$p_{i,n} = f(a_i s_n + b_i + e_{i,n}). \tag{1}$$

Here, a_i is a transfer coefficient between the Blinky and observed signals, b_i is a bias representing the background light (assumed to be time-invariant within the observation time), $e_{i,n}$ is the observation noise, and f is a function representing the saturation defined as

$$f(x) = \begin{cases} L & x > L \\ x & 0 \leq x \leq L \\ 0 & x < 0 \end{cases} \tag{2}$$

with a maximum brightness value L. As visualized in Fig. 2, (1) is represented in the matrix-vector form as

$$\mathbf{P} = \mathbf{f}(\mathbf{a}\mathbf{s}^\top + \mathbf{b}\mathbf{1}^\top + \mathbf{E}), \tag{3}$$

where $\mathbf{f}$ denotes the elementwise operation of f, the superscript $^\top$ denotes transpose of a vector or matrix, and

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & \cdots & p_{1,N} \\ \vdots & \ddots & \vdots \\ p_{I,1} & \cdots & p_{I,N} \end{bmatrix}, \tag{4}$$

$$\mathbf{s} = [s_1, \ldots, s_N]^\top, \tag{5}$$

$$\mathbf{a} = [a_1, \ldots, a_I]^\top, \tag{6}$$

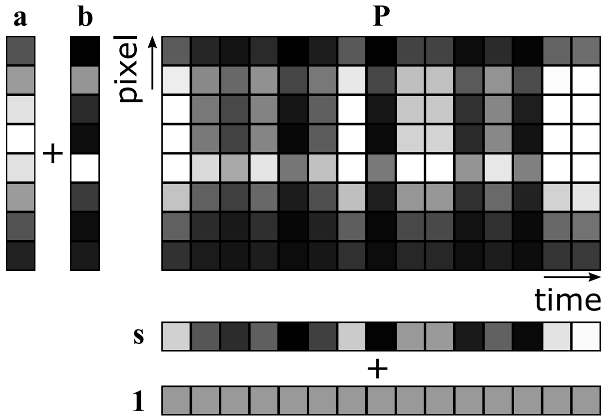


Fig. 2: Observation model. White pixels indicate high brightness.

$$\mathbf{b} = [b_1, \ldots, b_I]^\top, \tag{7}$$

$$\mathbf{E} = \begin{bmatrix} e_{1,1} & \cdots & e_{1,N} \\ \vdots & \ddots & \vdots \\ e_{I,1} & \cdots & e_{I,N} \end{bmatrix}. \tag{8}$$

Here, I and N are the numbers of pixel and time indices, respectively. Furthermore, (3) can be rewritten as

$$\mathbf{P} = \mathbf{f}(\mathbf{A}\mathbf{S} + \mathbf{E}), \tag{9}$$

where $\mathbf{A}$ and $\mathbf{S}$ are Rank-2 matrices defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{a} & \mathbf{b} \end{bmatrix}, \tag{10}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{s} & \mathbf{1} \end{bmatrix}^\top. \tag{11}$$

Since the light signals are generally represented by nonnegative values, the observation model (9) is similar to the nonnegative matrix factorization (NMF) [15] except for the existence of the saturation function $\mathbf{f}$.

III. Proposed Method

A. Calibration

First, we consider the calibration, i.e., the estimation of $\mathbf{a}$ and $\mathbf{b}$, which is required as the preliminary stage for estimation of the Blinky signal $\mathbf{s}$. Since these parameters are dependent on the target environment and generally difficult to know in advance, the calibration is performed by using the given reference signal $\mathbf{s}$. Here, it should be noted that the reference and observed signals are not synchronized because the Blinkies are not connected to the network. For their synchronization, the time difference l between the observed signal $\mathbf{P}$ and the unsynchronized reference signal, denoted by $\mathbf{s}'$, can be estimated by

$$l = \operatorname*{argmax}_m \sum_n s'_{n+m} \sum_{i \in R_n} p_{i,n}, \tag{12}$$

where R_n is the set of unsaturated pixel indices at time n, i.e., $i \in R_n \Leftrightarrow p_{i,n} < L$. Then, the reference signal can be synchronized to the observed signal as

$$s_{i,n} = s'_{i,n-m}. \tag{13}$$

Using the given synchronized reference and observed signals $\mathbf{s}$ and $\mathbf{P}$, the transfer coefficients $\mathbf{A}$ can be estimated by

$$\underset{\mathbf{A}}{\text{minimize}} \;\; J(\mathbf{A}) = \|\mathbf{M} \odot (\mathbf{A}\mathbf{S} - \mathbf{P})\|_F^2, \tag{14}$$

where $\odot$ denotes the Hadamard product (the elementwise product), $\|\cdot\|_F$ is the Frobenius norm of a matrix, and $\mathbf{M}$ is a binary matrix whose (i,n)th element $m_{i,n}$ is given by

$$m_{i,n} = \begin{cases} 0 & p_{i,n} = L \\ 1 & p_{i,n} < L \end{cases}. \tag{15}$$

Therefore, the objective function represents the sum of the squared error for the unsaturated pixels. Since the objective function $J(\mathbf{A})$ is a quadratic function with respect to $\mathbf{A}$, it can be minimized at

$$\begin{bmatrix} \hat{a}_i \\ \hat{b}_i \end{bmatrix} = (\mathbf{S}\mathrm{diag}(\mathbf{m}_{i,:})\mathbf{S}^\top)^{-1}(\mathbf{S}\mathrm{diag}(\mathbf{m}_{i,:})\mathbf{p}_{i,:}) \tag{16}$$

for each i, where $\hat{a}_i$ and $\hat{b}_i$ is the optimal value of a_i and b_i, respectively, $\mathbf{m}_{i,:} = [m_{i,1}, \ldots, m_{i,N}]^\top$, and $\mathbf{p}_{i,:} = [p_{i,1}, \ldots, p_{i,N}]^\top$.

B. Signal Restoration

Using the transfer coefficients $\mathbf{A}$, which is estimated by the above calibration method, we can estimated the unknown Blinky signal $\mathbf{s}$. This estimation problem is formulated as

$$\underset{\mathbf{s}}{\text{minimize}} \;\; K(\mathbf{s}) = \|\mathbf{M} \odot (\mathbf{a}\mathbf{s}^\top + \mathbf{b}\mathbf{1}^\top - \mathbf{P})\|_F^2, \tag{17}$$

whose objective function represents the sum of the squared error for the unsaturated pixels similarly as in the calibration method. Since the objective function $K(\mathbf{s})$ is a quadratic function with respect to $\mathbf{s}$, it can be minimized at

$$\hat{s}_n = (\mathbf{a}^\top \mathrm{diag}(\mathbf{m}_{:,n})\mathbf{a})^{-1}(\tilde{\mathbf{p}}_{:,n}^\top \mathrm{diag}(\mathbf{m}_{:,n})\mathbf{a}) \tag{18}$$

for each n, where $\hat{s}_n$ is the optimal value of s_n, $\mathbf{m}_{:,n} = [m_{1,n}, \ldots, m_{I,n}]^\top$, and $\tilde{\mathbf{p}}_{:,n} = [p_{1,n} - b_1, \ldots, p_{I,n} - b_I]^\top$.

IV. Experiments

Several experiments were conducted under simulated and real environments to evaluate and compare the estimation accuracy of the proposed method and several other methods.

A. Simulated Environment

Here, we evaluated the accuracy of the calibration and signal reconstruction under the ideally simulated environment where the observed signal was given exactly by (1). The pixel size of the observed signal was 32×32 ($I = 1024$), and the number of samples was $N = 150$, which corresponds to five seconds of 30-fps video. The true transfer coefficient $\mathbf{a}$ was a two-dimensional Gaussian function as shown in Fig. 3(a), and the true bias $\mathbf{b}$ was a random vector each of whose element was sampled independently from the uniform distribution in the interval $[0, 0.4]$ as shown in Fig. 4(a).

TABLE I: Comparison of SNR (dB) for calibration in simulated environment.

Parameter	Transfer coefficient $\mathbf{a}$	Bias $\mathbf{b}$
Conventional [9]	7.467	5.882
Proposed	27.398	25.886

First, the accuracy of the calibration method was evaluated. The reference signal, denoted by $\mathbf{s}_{\mathrm{ref}}$, was a sawtooth wave having two periods with a period of 75 samples. The observation noise $e_{i,n}$ was sampled independently from the Gaussian distribution with mean 0 and variance 0.05^2. Here, the following two methods were compared:

- Conventional [9]: The estimated transfer coefficient and bias were given by

$$\hat{a}_i = \hat{a} = \frac{1}{I}\left(\max_n \sum_i p_{i,n} - \min_n \sum_i p_{i,n}\right), \tag{19}$$

$$\hat{b}_i = \hat{b} = \frac{1}{I}\min_n \sum_i p_{i,n}. \tag{20}$$

  Note that this method does not consider the saturation and that $\hat{a}_i$ and $\hat{b}_i$ are invariant for i.
- Proposed.

The Signal-to-Noise Ratio (SNR) defined as

$$\mathrm{SNR}(\hat{\mathbf{a}}, \mathbf{a}) = 10\log_{10}\frac{\sum_i a_i^2}{\sum_i(\hat{a}_i - a_i)^2} \tag{21}$$

for $\mathbf{a}$ (and also for $\mathbf{b}$) was used as the evaluation criterion.

The SNRs for Conventional [9] and Proposed are shown in Table I. One can see that Proposed achieved much higher SNRs than Conventional [9] for both $\mathbf{a}$ and $\mathbf{b}$. In addition, the calibration results for $\mathbf{a}$ and $\mathbf{b}$ are plotted in Figs. 3 and 4, respectively. These figures also indicate that the transfer coefficient $\mathbf{a}$ and bias $\mathbf{b}$ were estimated with high accuracy in Proposed.

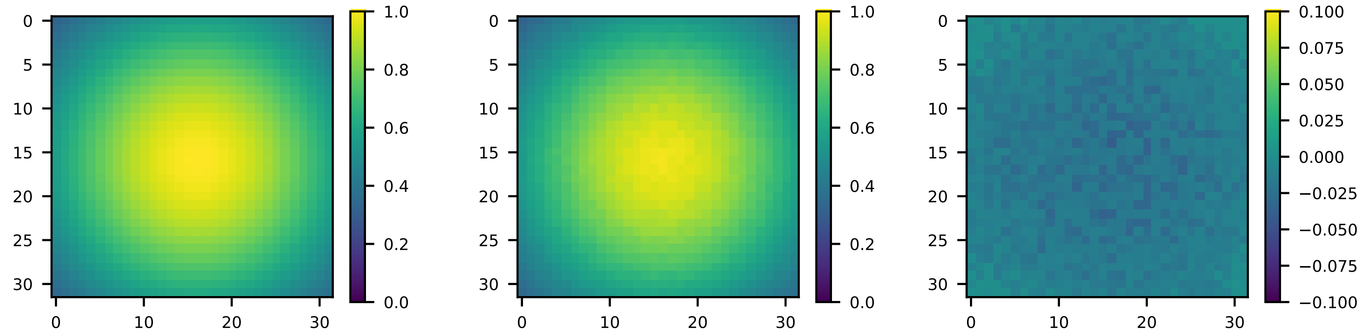
Next, the estimation accuracy for the Blinky signal was evaluated. In the signal restoration, the following three signals were used as the Blinky signals $\mathbf{s}$.

- Random: each s_n was sampled independently from the Gaussian distribution with mean 0.5 and variance 0.5^2.
- Sinusoids:

$$s_n = \frac{1}{4}\left(\sin\left(\frac{2\pi}{5}\frac{n}{F}\right) + \sin\left(\frac{4\pi}{5}\frac{n}{F}\right)\right) + \frac{1}{2}, \tag{22}$$

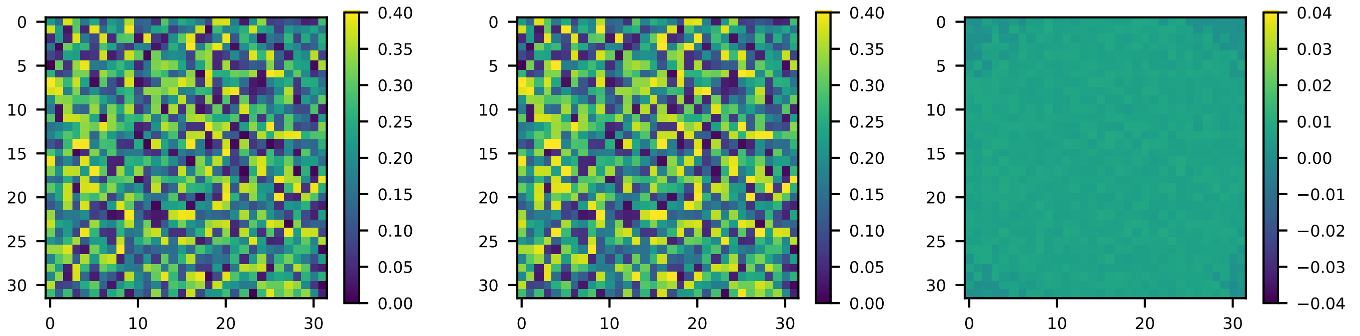
  with $F = 30$.
- Music: a violin sound from SMILE2004 [16], which is converted to the Blinky signal in accordance with [7].

The observation noise $e_{i,n}$ was added in the same manner as in the calibration. To evaluate the estimation accuracy of the Blinky signal, we used $\mathrm{SNR}(\hat{\mathbf{s}}, \mathbf{s})$. Here, the following four methods were compared.



(a) True value (b) Estimated value (c) Difference between true and estimated values

Fig. 3: Calibration results of transfer coefficient $\mathbf{a}$ in simulated environment.



(a) True value (b) Estimated value (c) Difference between true and estimated values

Fig. 4: Calibration results of transfer bias $\mathbf{b}$ in simulated environment.

- Pixel mean: The average of the observed signals for all pixels was used to estimate the Blinky signal.

$$\hat{s}_n = \frac{\gamma}{I} \sum_i p_{i,n} + \delta, \tag{23}$$

where γ and δ were given by

$$\gamma = \frac{\sum_n ((\sum_i p_{i,n}/I - \bar{p})(s_n - \bar{s}))}{\sum_n (\sum_i p_{i,n}/I - \bar{p})^2}, \tag{24}$$

$$\delta = \bar{s} - \gamma \bar{p}, \tag{25}$$

with $\bar{s} = \sum_n s_n/N$ and $\bar{p} = \sum_{i,n} p_{i,n}/(IN)$ so that the SNR was maximized.

- Best pixel: The one pixel achieving the highest SNR was used to estimate the Blinky signal.

$$\hat{\mathbf{s}} = \frac{\mathbf{p}_{\hat{i},:} - b_{\hat{i}}\mathbf{1}}{a_{\hat{i}}}, \tag{26}$$

$$\hat{i} = \underset{i}{\mathrm{argmax}}\ \mathrm{SNR}\left(\frac{\mathbf{p}_{i,:} - b_i\mathbf{1}}{a_i}, \mathbf{s}\right). \tag{27}$$

Here, $\mathbf{a}$ and $\mathbf{b}$ were not the estimated values but the true values.

- Conventional [9]: The estimated Blinky signal was given by

$$\hat{s}_n = \frac{\hat{a}}{I} \sum_i p_{i,n} + \hat{b}, \tag{28}$$

where $\hat{a}$ and $\hat{b}$ were obtained from (19) and (20), respectively.

- Proposed.

TABLE II: Comparison of SNR (dB) for signal restoration in simulated environment.

Blinky signal	Random	Sinusoids	Music
Pixel mean	20.422	17.900	21.884
Best pixel	23.425	23.013	26.585
Conventional [9]	10.650	8.996	7.958
Proposed	31.063	29.147	28.813

Note that Pixel mean and Best pixel not only required the true coefficients $\mathbf{a}$ and $\mathbf{b}$, but also used the true Blinky signal $\mathbf{s}$ to determine the optimum parameters of γ, δ and $\hat{i}$ as oracle information. Then, they cannot be used directly in practical situations.

The result of the SNR comparison was shown in Table II, where Proposed achieved the highest SNRs for all Blinky signals. This result showed the effectiveness of the proposed method for integrating multiple pixels with consideration of

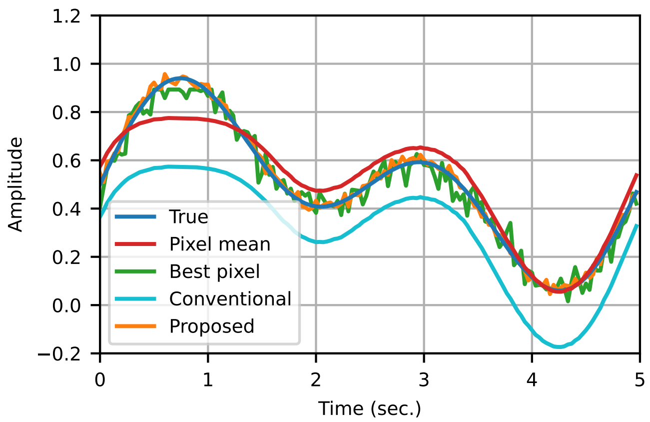


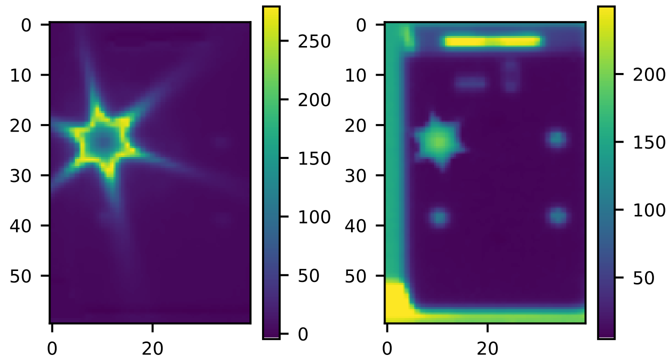
Fig. 5: True and estimated Blinky signals in simulated environment.

Fig. 6: Experimental environment. Pixels enclosed by red box are used in calibration and signal restoration.

the saturation. For further investigation, the true and estimated Blinky signals for Sinusoids are plotted in Fig 5. One can see that the amplitude ranges of Pixel mean and Conventional were narrower than the other two methods. This was because the saturated observed signals were used without any compensation in Pixel mean and Conventional. Best pixel had a sufficient amplitude range; however, it was affected significantly by the observation noise. Proposed was able to suppress the noise effect while keeping the amplitude range by integrating the nonsaturated pixels and times.

B. Real Environment

Experiments were also conducted in a real environment. A Blinky was observed in gray scale with an industrial camera whose frame rate was 30 Hz as shown in Fig. 6. The 40×60 pixels ($I = 2400$) around the Blinky (i.e., pixels enclosed by red box in Fig. 6) were used in the calibration and signal restoration. The reference signal $\mathbf{s}_{\mathrm{ref}}$ in the calibration was the sawtooth wave repeated three times, and the Blinky signal $\mathbf{s}$ in the signal restoration was Sinusoids repeated twice.



(a) Estimated transfer coefficient $\hat{\mathbf{a}}$ (b) Estimated bias $\hat{\mathbf{b}}$

Fig. 7: Calibulation results in real environment.

TABLE III: Comparison of SNR (dB) for signal restoration in real environment.

Blinky signal	Sinusoids
Pixel mean	23.707
Best pixel	31.114
Conventional [9]	10.343
Proposed	23.781

Since true transfer coefficient and bias were unknown in the real environment, only the estimation accuracy for the Blinky signal was evaluated. The same four methods as in the previous experiment were compared; however, unlike in the previous experiment, transfer $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ estimated by Proposed were also used in Best pixel.

First, Fig. 7 shows the calibration results. Since true $\mathbf{a}$ and $\mathbf{b}$ were unknown, the quantitative evaluation is difficult in this case. However, one can see the large transfer coefficient around the LED of the Blinky, which means the transfer coefficient was considered to be estimated accurately to some degree. On the other hand, the estimated transfer coefficient was lower in the central area of the LED than in its peripheral area, and the estimated bias had large values in the central area of the LED although it had to be independent from the Blinky signal. These unexpected results were considered to be estimation errors caused by the strong saturation in the central area of the LED; the observed signals in these pixels were almost always saturated and therefore not applicable in the calibration and the signal restoration.

Next, the SNRs for the signal restoration are shown in Table III. The SNR of Proposed was lower than that in the simulated environment, which was considered to be caused by the calibration error described in the previous paragraph. However, Proposed still achieved much higher SNR than Conventional [9], which indicated the validity of Proposed among the methods without oracle information. Finally, the true and estimated Blinky signals are plotted in Fig. 8. Similar tendencies as in the simulated environment are also seen in this condition, and these results indicated that Proposed was still able to restore the Blinky signal in the real environment

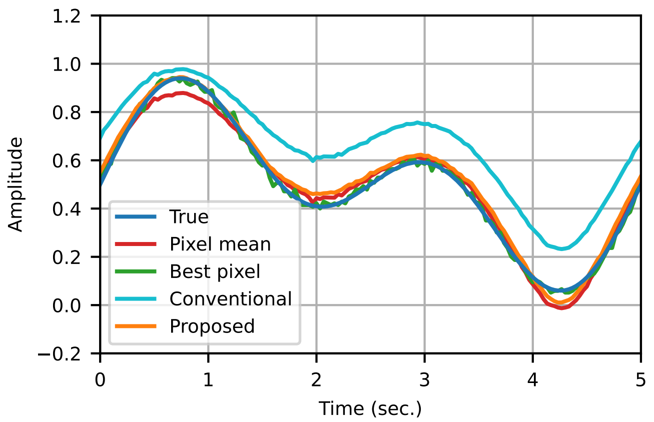


Fig. 8: True and estimated Blinky signals in real environment.

even though its SNR was lower to some extent than that in the simulated environment.

V. Conclusion

We proposed methods for calibrating the transfer coefficients and restoring the Blinkiy signals considering the saturation effect of the observed signals. By focusing the fact that a single Blinky signal can be observed at multiple pixels for different brightness, we formulated the calibration and signal restoration problems as the least-squares problems for nonsaturated pixels and times. Experiments under numerical and real environments demonstrated the validity of the proposed method. In future work, we will consider a more sophisticated calibration and signal restoration methods incorporating the NMF frameworks.

Acknowledgement

This work was supported by JST CREST Grant Number JPMJCR19A3 and JSPS KAKENHI Grant Number JP20H00613.

References

[1] M. Chen, Z. Liu, L.-W. He, P. Chou, and Z. Zhang, "Energy-based position estimation of microphones and speakers for ad hoc microphone arrays," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2007, pp. 22–25.

[2] M. S. Brandstein and H. F. Silverman, "A practical methodology for speech source localization with microphone arrays," *Computer Speech & Language*, vol. 11, no. 2, pp. 91–126, 1997.

[3] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2011, pp. 189–192.

[4] N. Ono, "Fast stereo independent vector analysis and its implementation on mobile phone," in *Proc. International Workshop on Acoustic Signal Enhancement*, 2012, pp. 1–4.

[5] A. Temko, R. Malkin, C. Zieger, D. Macho, C. Nadeu, and M. Omologo, "CLEAR evaluation of acoustic event detection and classification systems," in *Proc. International Evaluation Workshop on Classification of Events, Activities and Relationships*, 2007, pp. 311–322.

[6] D. Giannoulis, E. Benetos, D. Stowell, M. Rossignol, M. Lagrange, and M. D. Plumbley, "Detection and classification of acoustic scenes and events: An IEEE AASP challenge," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2013, pp. 1–4.

[7] R. Scheibler and N. Ono, "Blinkies: Open source sound-to-light conversion sensors for large-scale acoustic sensing and applications," *IEEE Access*, vol. 8, pp. 67603–67616, 2020.

[8] R. Scheibler and N. Ono, "Multi-modal blind source separation with microphones and blinkies," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2019, pp. 366–370.

[9] D. Horiike, R. Scheibler, Y. Kinoshita, Y. Wakabayashi, and N. Ono, "Energy-based multiple source localization with Blinkies," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 2020, pp. 443–448.

[10] K. Ishii, Y. Kinoshita, Y. Wakabayashi, and N. Ono, "Real-time pitch visualization with "Blinky" sound-to-light conversion device," *Journal of Signal Processing*, vol. 25, no. 6, pp. 213–220, 2021.

[11] Y. Kinoshita and N. Ono, "End-to-end training for acoustic scene analysis with distributed sound-to-light conversion devices," in *Proc. European Signal Processing Conference*, 2021, pp. 1010–1014.

[12] K. Siedenburg, M. Kowalski, and M. Dörfler, "Audio declipping with social sparsity," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014, pp. 1577–1581.

[13] Ç. Bilen, A. Ozerov, and P. Pérez, "Solving time-domain audio inverse problems using nonnegative tensor factorization," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5604–5617, 2018.

[14] P. Záviška, P. Rajmic, A. Ozerov, and L. Rencker, "A survey and an extensive evaluation of popular audio declipping methods," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 1, pp. 5–24, 2020.

[15] A. Cichocki, R. Zdunek, and S. Amari, "Nonnegative matrix and tensor factorization [lecture notes]," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 142–145, 2008.

[16] K. Kawai, K. Fujimoto, T. Iwase, H. Yasuoka, T. Sakuma, and Y. Hidaka, "Development of a sound source database for environmental/architectural acoustics: Introduction of SMILE 2004 (Sound Material in Living Environment 2004)," in *Proc. International Congress of Acoustics*, 2004, pp. 1561–1564.