

# 俯瞰映像合成システム Fairyview<sup>®</sup>のユーザビリティ評価

皆川純<sup>1</sup> 岡原浩平<sup>1</sup> 山崎賢人<sup>1</sup> 深澤司<sup>1</sup>

**概要:** 映像監視業務において、複数のカメラ映像を並べて表示すると、カメラ台数の増加に比例して業務負荷が増加する。さらに、各映像が独立しているため、空間把握が難しい。そこで、複数カメラの映像を合成し、1枚の俯瞰映像にすることで、映像監視業務の負荷を減らすだけでなく、監視エリアの状況把握を容易にするシステムを開発した。本論文では、俯瞰映像合成システムの概要とともに、実証実験環境を構築したことについて述べる。さらに、実証実験環境にて生成した俯瞰映像に対して、アンケートによるユーザビリティ評価を実施し、有用性を確認した。

**キーワード:** 映像合成, マルチカメラ, 映像監視システム

## Usability Evaluation for an Aerial-view Video Synthesis System

JUN MINAGAWA<sup>†1</sup> KOHEI OKAHARA<sup>†1</sup>  
KENTO YAMAZAKI<sup>†1</sup> TSUKASA FUKASAWA<sup>†1</sup>

**Abstract:** This paper describes the concept and building a demonstration environment of Fairyview<sup>®</sup>. The more increase conventional video surveillance system has cameras, the more difficult operators understand situations of the whole area due to a lot of camera images displayed separately. To address this challenge, we propose an aerial-view video synthesis system named Fairyview<sup>®</sup>, which process each camera image into a top-view stitched image. This system enables operators to grasp surroundings quickly through the stitched image. Additionally, usefulness of proposed system was shown by a usability evaluation using questionnaires.

**Keywords:** Image Stitching, Multi-Camera, Video Surveillance System

### 1. はじめに

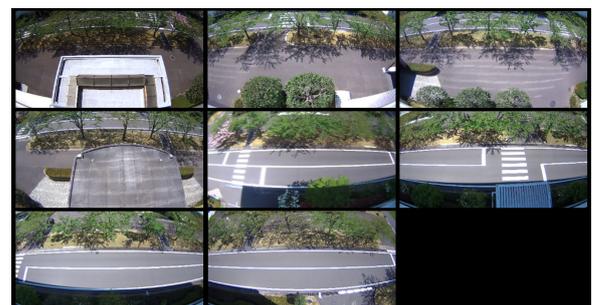
安心・安全に対する関心の高まりにつれて監視カメラの設置台数は年々増加しており、映像監視業務の負荷は増加の一途を辿る。図 1(a)に示すとおり、従来の映像監視システムでは、同時に複数のカメラ映像を監視者に提供するために、監視モニターに各カメラの映像を並べて表示していた。しかし、この表示方法は同時にすべてのカメラ映像を監視者に提供できる反面、画面上に表示されるカメラ映像の位置関係を即座に判断し、監視エリア全体の状況を把握するのが難しい。また、従来のシステムでは注視している監視対象の動線を監視者が確認する場合、分割表示されているカメラ映像の中から、注視している人物を同定し、その人物の移動経路を頭の中でイメージしなければならない。

そこで我々は、監視エリアの空間把握を容易にするために、複数の監視カメラの映像から、エリア全体を俯瞰する俯瞰映像合成システム Fairyview<sup>®</sup>を開発してきた[1]。Fairyview<sup>®</sup>は、各カメラ映像が1枚の連続した映像として監視者に提示されるため、人物の位置や動線を把握する場合において効果を発揮する(図 1(b))。さらに、我々は産業的な側面から様々な機能拡張を試行してきた。例えば、文献[2]では長時間設置したカメラの自重や飛来物の衝突によりカメラの位置姿勢が変化したこと起因する合成映像

のズレを自動修正する手法を提案した。

こうした研究開発をもとに Fairyview<sup>®</sup>は三菱電機より製品化され、2020年には2020 R&D 100 Awards Finalistに選出された[3]。さらに、Fairyview<sup>®</sup>は SECURITY SHOW 2016に出展するなど、産業分野におけるデジタルコンテンツ生成技術の浸透にも貢献してきた。

本論文では、三菱電機イベントスクエア METoA Ginzaでの展示時にとったアンケートをもとにユーザビリティを評価し、Fairyview<sup>®</sup>の有用性を確認したことについて述べる。



(a) 従来システムの映像の表示 (整列表示) 例



(b) 俯瞰映像の表示例

図 1 表示例

<sup>1</sup> 三菱電機株式会社  
Mitsubishi Electric Corporation

## 2. 関連研究

本章では、Fairyview®と同様に多視点画像を合成する製品・サービスの例を挙げる。

Google Street View では静止画像をもとに生成した全周囲画像を表示するサービスを提供している。

一方、リアルタイムに映像を合成し表示するものも多い。日産自動車では、駐車時の周辺監視を支援することを目的に、自動車に搭載したカメラ画像を用いて自動車を俯瞰できるアラウンドビューモニターを開発した[4]。各自動車メーカーも同様のシステムを製品化しており、自動車の運転において、俯瞰合成映像を利用することは一般化している。

リアルタイムの映像を合成する場合、静止画を合成するよりも処理時間の制約が大きい。その対策として、カメラの校正結果から、事前に定義した画像合成面にカメラ画像をマッピングする手法が多い。定義する画像合成面は目的にあった形状にする必要がある。アラウンドビューモニターは画像合成面を地面（平面）と定義し、広範囲を俯瞰する映像を表示する。沖電気工業のフライングビューは、お椀形状の画像合成面を使用することで、視点の自由度が高い周辺監視を実現している[5]。

映像合成により周辺監視の効率は向上するため、一般車両向けだけでなく、建設機械向けに利用されるケースが増加している[6]。

このように、俯瞰映像合成技術は自動車に搭載したカメラで応用されるケースが多かったが、近年では、高速道路料金所など、高所につけたカメラを用いたケースも登場している。さらに、工場内にカメラを取り付け、製造現場の業務改善のため映像解析と連携するなど、高度な利用方法も提案されている[7]。

空間を把握しやすい俯瞰映像合成技術のニーズは高まっている。3章では、監視業務における視認性の向上に適したシステムを構築したことについて述べる。

## 3. 俯瞰映像合成システム Fairyview®

本章では、俯瞰映像合成システム Fairyview®のコアとなる映像合成手法について述べるとともに、実装した実証実験環境について説明する。

### 3.1 映像合成手法

1章で述べたとおり、Fairyview®は広い監視エリアの状況把握を目的としているため、地面を映像合成面にする。そのため、高所に複数のカメラを設置し、それらの映像をもとに合成する必要がある。

上記で定義した映像合成面で各カメラの映像を合成するためには、取得した画像のレンズ歪みを補正後、画像合成面に対する各カメラの外部パラメータを推定し、その推定した外部パラメータをもとに画像合成面に映像を投影する必要がある。このとき、隣接するカメラ映像が重畳する領域では $\alpha$ ブレンドにより違和感のない1枚の合成映像を

生成する。

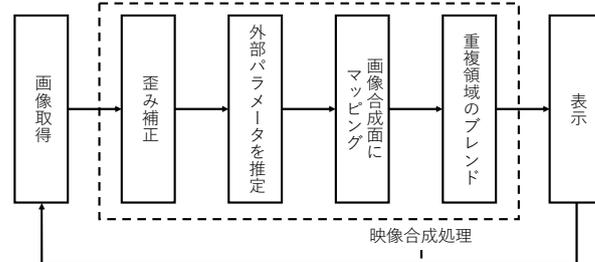
しかし、Fairyview®は監視目的で使用するために、映像をリアルタイムに表示する必要がある。リアルタイムに複数のカメラの映像を俯瞰合成するために、1フレームごとに前述の処理をしては、フレームレートが下がる。これらの処理をリアルタイムに実現するために、カメラの位置姿勢が不変であることを前提に、事前に外部パラメータを推定するとともに、任意のカメラ映像の各画素と俯瞰映像の各画素との対応関係をあらかじめ計算した画素対応マップを作成しておく。そうすることにより、毎フレームの合成処理は画素対応マップを参照するだけで俯瞰映像を生成することができる（図2）。

前述したとおり、Fairyview®を構築する際には、外部パラメータのカメラキャリブレーションを事前に実施する。具体的には、実世界の監視エリア上にGCP(Ground Control Point)を配置し、カメラ映像中のGCPから外部パラメータを推定する。

### 3.2 システム構成

3.1節で述べた映像合成手法を元に、三菱電機情報技術総合研究所内にFairyview®の実証実験環境を構築した（図3）。実証実験環境における監視エリアは、22m×80mの屋外である。監視エリアを撮影するカメラは図3に示すように監視エリアの周囲にある建物の壁面に等間隔に8台配置した。設置したカメラは、一般的な映像監視システムに用いられるネットワークカメラを使用した。各カメラは、俯瞰映像を生成するためのサーバ（映像合成サーバ）とネットワークで接続しており、映像（解像度フルHD(1920px×

一般的な処理フロー



Fairyview®の処理フロー

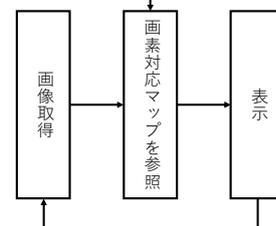


図2 処理フロー

表1 映像合成サーバの仕様

OS	CentOS 7.6
CPU	Intel Xeon E5-1620 v2 3.90GHz
Memory	16GB
GPU	Nvidia Quadro M5000 8GB

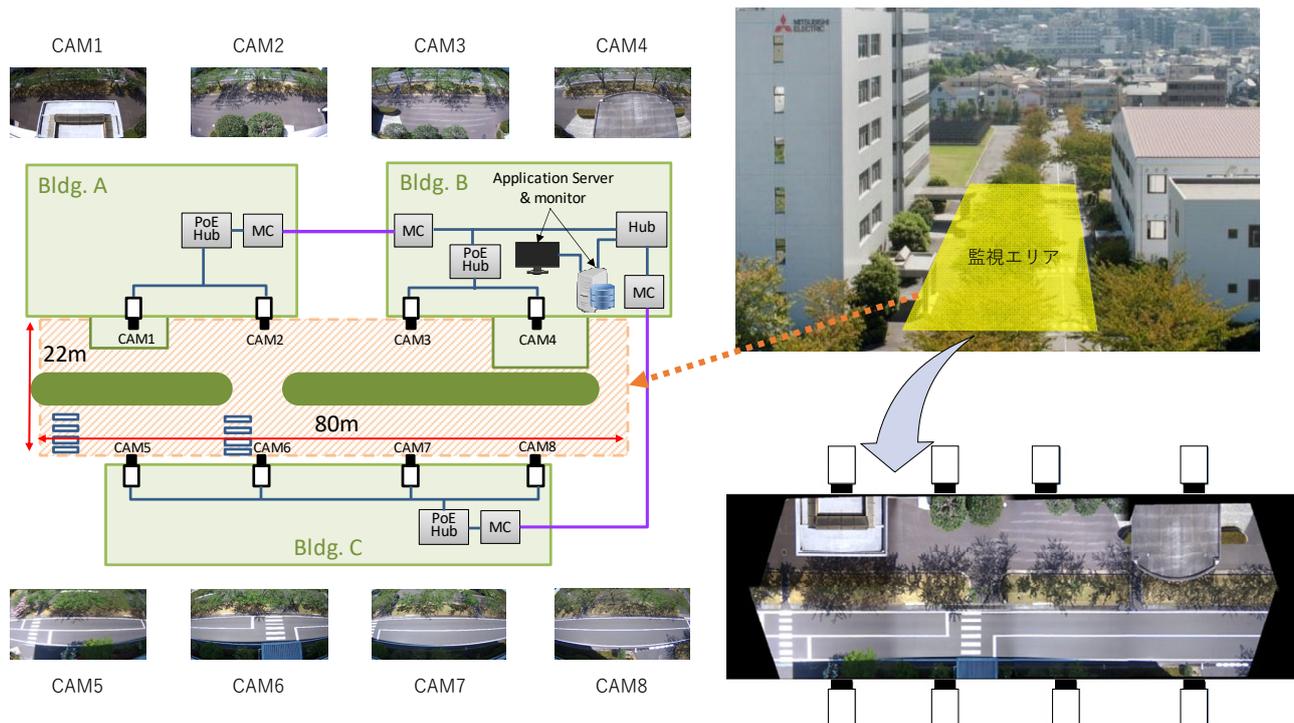


図 3 Fairyview®の構築例

1080px), フレームレート 30) を, 映像合成サーバへ配信する。

図 3 に示した実証実験環境における映像合成サーバは, 3.1 節で述べた映像合成処理を行う。本実証実験環境では, 8 枚のカメラ映像から生成したフル HD 解像度の俯瞰映像を, 映像合成サーバに接続したモニターで表示している。このモニターにフレームレート 30, 遅延 2s 以内で表示することができる。ただし, 表 1 に示した仕様は実証実験環境における映像合成サーバであり, 製品版とは異なる。

#### 4. Fairyview®のアンケート調査

本章では, 多くの人に Fairyview®が提供する俯瞰的な表示方法が有用であるかどうかを検証することを目的に実施したアンケートについて述べる。

##### 4.1 アンケート方法と内容

我々は, METoA Ginza にて, 2021 年 9 月から 2022 年 1 月の期間に Fairyview®を出展した。イベント会場では, 3.2 節に示した実証実験環境のデモ映像とともに, Fairyview®の有効性を示すために, 従来の監視システムの映像の表示と Fairyview®による俯瞰映像の表示例の比較画像を掲示し, アンケートを取った。図 4 に示すのは, イベント会場の様子とアンケート記入台である。アンケートでは, 従来の監視システムの表示方法と Fairyview®による俯瞰映像を比較し, 次の 3 つの設問に対して「そう思う」「ややそう思う」「どちらとも言えない」「あまりそう思わない」「そう思わない」の 5 段階で回答してもらった。



図 4 METoA Ginza での展示の様子

- (a) 全体の様子が分かりやすい
- (b) 人物が地図上のどこにいるか分かりやすい
- (c) 人物が何をしているか分かりやすい

年齢別の回答者数は 19 歳以下が 27 人, 20 代が 52 人, 30 代が 71 人, 40 代が 41 人, 50 代以上が 21 人であり, 未回答 3 人を含む計 215 人から回答を得た。

##### 4.2 アンケート結果

アンケート結果の設問ごとの回答結果を図 5 に示す。

設問(a)「全体の様子が分かりやすい」において, そう思う・ややそう思うと回答した人は, 従来映像は 147 人, 俯瞰映像は 203 人だった。俯瞰映像に対して全体の様子が分かりにくいと感じた人はごく少数であった。また, 設問(b)「人物が地図上のどこにいるか分かりやすい」において, そう思う・ややそう思うと回答した人は, 従来映像は 145 人, 俯瞰映像は 200 人であった。本設問においても分かり

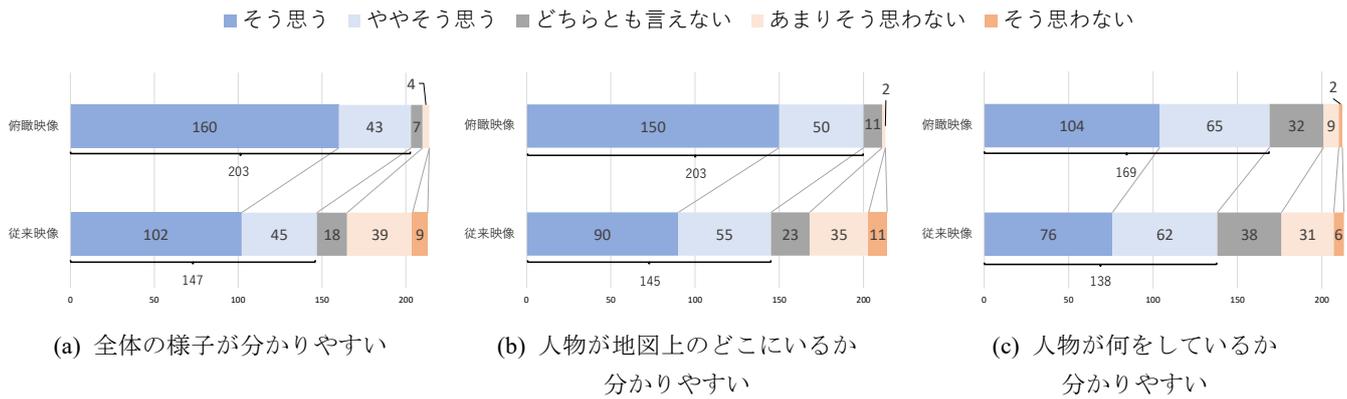


図 5 アンケート結果

にくいと感じた人は極めて少数であった。

一方で、設問(c)「人物が何をしているか分かりやすい」については、そう思う・ややそう思うと回答した人は、従来映像は 138 人、俯瞰映像は 169 人だった。設問(a)(b)では、俯瞰映像によって分かりやすいと感じた人が 50 人以上増えていた。しかし、本設問にて分かりやすいと回答した人は約 30 人しか増えなかった。

前記の設問の他に、アンケート内の自由記述欄に対しても様々な意見が寄せられた。例えば、「道案内をするときの見せ方につかえそう」「観光地の混雑具合を把握するのに役立つそう」などの意見があった。また、「スポーツ観戦への適用ができそう」というコメントもあった。一方で、「何を目的にするかによって、見せ方の良し悪しは変わると思う」といった意見もあった。

### 4.3 考察

アンケート結果から、監視エリアを俯瞰的な合成映像として表示する方法は、監視エリア全体の様子が分かりやすい点と、人物が地図上のどこにいるか分かりやすい点において有用であることが示された。ゆえに、1 章にて説明したシステムのユースケース「不審者の動線（特定のエリアをうろついている等）を判断でき、その人物の現在位置を特定できる」に対して有効であると判断できる。

一方で、人物が何をしているかどうかを判断する点において、俯瞰的な見せ方のほうが分かりやすいと感じた人は一定数いたが、そう思わない人も一定数いた。おそらく、人物の手元や顔といった部分が、俯瞰視点だと分かりにくいと感じた人がいたのではないかと予想する。例えば、人物の顔や服装、ジェスチャ（手の動きなど）から不審者と判断するのは難しいと思われる。また、俯瞰映像による人物の動線から不審者を判断し、その位置を特定できたとしても、その不審者の詳細な情報を得にくいと考えられる。アンケート内のコメントにあったように、俯瞰映像による見せ方は、目的によってメリットとデメリットが分かれることが明らかになった。

上記の問題の解決策として、俯瞰映像によって分かりにくくなった情報を補助するための機能が有効であると考え

る。例えば、俯瞰映像上の人物に監視者がフォーカスしたら、その人物を捉えているカメラの映像を自動選択し、映像中の当該人物の画像領域を抜き出して、俯瞰映像上にポップアップ表示することが考えられる。また、より詳細な人物の映像を取得する方法として、PTZ カメラと連携した表示方法も有効であると考えられる。

## 5. まとめ

本論文では、監視業務の視認性向上を目的として、監視エリアを真上から見た時の映像を生成し、監視者へ提供することで、監視エリアを一目で確認できる俯瞰映像合成システム Fairyview<sup>®</sup>について述べた。また、一般人向けのイベントに Fairyview<sup>®</sup>を出展し、俯瞰的な見せ方による映像監視方法の有効性をアンケート調査した。アンケートの結果、監視エリア全体の様子と人物の位置を把握することにおいて、俯瞰映像は有効であると示された。

一方で、人物の詳細を把握することにおいては、改善が必要であると分かった。今後は、Fairyview<sup>®</sup>の機能改善として、人物の詳細を提示するための補助機能の検討を進めていく。

## 参考文献

- [1] 菊池他：俯瞰映像合成技術“Fairyview”，三菱電機技報，Vol. 94, No. 9, pp. 512-515(2020).
- [2] Minagawa, J., et al.: A camera recalibration method for a topview surveillance system based on relative camera pose and structural similarity, 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance, IEEE, pp. 1-8 (2019).
- [3] R&D World: Finalists for 2020 R&D 100 Awards are unveiled, <https://www.rdworldonline.com/finalists-for-2020-rd-100-awards-a-re-unveiled/>
- [4] 小林他：フライングビュー：人の目と AI の目で見たいものを見逃さないリアルタイムリモートモニタリングシステム, Oki テクニカルレビュー, Vol. 88, No. 2, pp. 6-9 (2021).
- [5] 酒井：世界初アラウンドビューモニター，自動車技術= Journal of Society of Automotive Engineers of Japan, Vol. 62, No. 3, pp. 100-101 (2008).
- [6] 佐藤他：無人化施工用俯瞰映像提示システムの開発，日本機械学会論文集，Vol. 81, No. 823, pp. 14-31 (2015).
- [7] 日立産業制御ソリューションズ：映像処理ソリューション，[https://info.hitachi-ics.co.jp/product/image\\_sol/](https://info.hitachi-ics.co.jp/product/image_sol/)