

時空間の異なる 100m 走選手間での AR 仮想対決

八木 賢太郎*¹ 長谷川 邦洋*¹ 齋藤 英雄*¹

Virtual 100m Race with Different Time and Space Runners

Kentaro Yagi*¹ Kunihiro Hasegawa*¹ and Hideo Saito*¹

Abstract – 本論文では陸上競技の 100m 走を対象に、時空間の異なる選手間による仮想対決映像生成手法を提案する。本手法はオリンピックなどの TV・インターネット放映映像に対して新たな映像コンテンツとしての利用を想定しているため、カメラ位置は映像毎に異なり、かつズーム・パンを行うシーンを対象としている。異なる映像間において選手の投影位置を正しく求めるためには、各フレームでカメラが陸上トラックの何 m 地点を撮影しているかを求める必要がある。本手法ではトラック直線の連続性を利用し、各映像における 100m トラックのパノラマ画像を生成することで撮影領域の推定を行う。また生成したパノラマ画像間での平面射影行列を用いることで映像間での対応付けを行う。実験では実際に異なる時空間で撮影されたレース映像を用いて選手間の仮想対決映像を生成する。

Keywords : AR, Virtual Race, Image Stitching, Sports, Track and Field

1 はじめに

画像処理技術を用いて、スポーツ放映の新たなコンテンツを生み出そうという試みはこれまで多くなされて来た。その例としてスタジアムに搭載された複数カメラにより自由視点映像を生成する技術 [1] や、テニスボールを複数カメラでトラッキングし、ボールの軌跡を可視化したりインアウト判定を行うシステムが挙げられる [2]。

本論文では陸上の 100m 走を対象とし、異なる時空間で撮影された映像中の選手による仮想対決映像生成システムを提案する。本手法はカメラ位置が映像毎に異なり、かつズーム・パンを行うシーンを対象としている。また過去の映像中の選手との比較を想定しているため、上で示したシステムのようにマルチカメラを用いず単一視点から撮影された映像を用いる。一般的に移動カメラから撮影された映像上に物体を正しく重ねるためには SLAM や Structure from Motion 等を用いて、背景の 3 次元形状とカメラ位置を推定する必要がある [3][4]。しかしながら本論文が対象とする 100m 走映像中にはテクスチャが似通った画像が多く含まれるため特徴量のマッチングが難しく、またズームを行うためフレーム毎の内部パラメータは異なる。従って特徴量ベースの 3 次元形状復元手法を用いることはできない。

背景の 3 次元復元を行わずスポーツシーンでの仮想対決を実現した例として NHK 放送技術研究所による仮想対決カメラシステムが挙げられる [5]。本システムはスキー競技を対象とし、移動カメラによって撮影さ

れた映像上に異なるレースの選手を投影する。しかしながら、本システムは入力映像が全て、同じ位置から同じ軌跡を辿り撮影されたことを前提としており、用いられたカメラや軌跡が異なるシーンでは利用できない。

こうした研究に対して我々は、時空間が異なり、かつ撮影に用いられるカメラや移動軌跡が異なるシーンにおいても使用可能な仮想対決映像生成手法を提案する。

2 提案手法

ここでは提案手法の具体的な処理について述べる。図 1 にアルゴリズムの概要を示す。あるレース中の選手を別の映像に投影する場合、まず各フレームにおいて対象となる選手が何 m 地点を走行しているかを求める必要がある。しかしながら、陸上トラック上に配置された距離マーカーを映像から確認することは難しく、各フレームから選手の走行位置を求めることは不可能である。そのため本手法ではまず画像列からパノラマ画像を生成し、100m トラック全体の俯瞰画像を生成することで各フレーム中の選手の走行位置を求める。その後生成されたパノラマ画像中のスタートライン・ゴールライン上の角 4 点を用いて異なる映像間での斜射変換行列を求め選手の投影を行う。

2.1 パノラマ画像の生成

一般的に画像群からパノラマ画像を生成する際には SIFT[6] や AKAZE[7] のような局所特徴量を用いた画像のマッチング手法が用いられる [8]。しかし、提案手法で対象とするような陸上トラックは単調なテクスチャを持つため、検出される特徴点は似通った特徴量を持ち、画像をマッチングする際に誤対応が発生し易いという問題がある。

*¹慶應大学

*¹Keio University

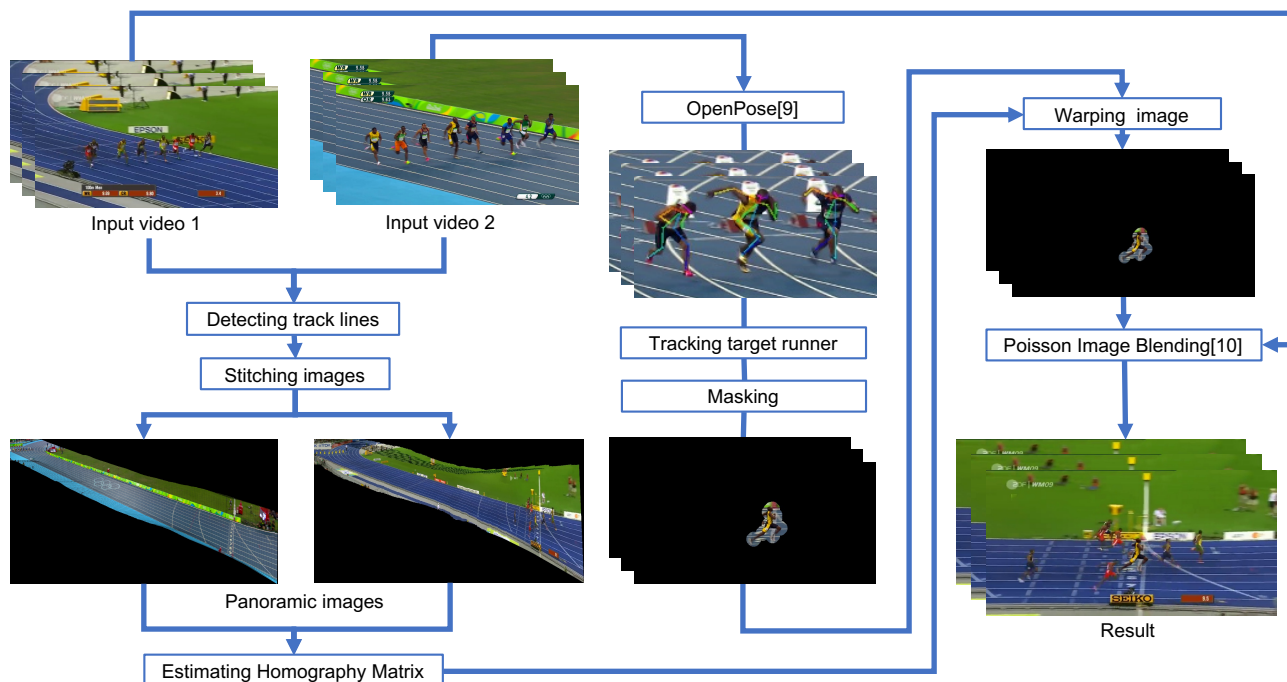


図1 提案手法の流れ

そこで提案手法では、トラックの直線情報を用いた画像のマッチング手法を提案する。直線情報を用いることで局所特徴量を用いたマッチング手法よりも頑健にパノラマ画像を生成することができる。

2.1.1 直線検出

ここでの目的は画像間のマッチングを行うために、選手の走行方向に伸びるトラック上の直線を検出し、かつ各直線のトラッキングを行うことである。

図2に直線検出処理の流れを示す。まず1フレーム目において各直線上の2点をクリックすることで検出したい直線の初期位置を決定する。2フレーム目以降ではRANSACと最小二乗法を用いて直線の検出を行う。まず縦方向の微分フィルタを適用することで画像水平方向のエッジを検出する。一般的に100mのレースではカメラは選手の走行方向と平行に移動を行うため、検出したいトラック上の直線は水平方向のエッジに含まれていると考えられる。その後、直前のフレームで検出された各直線の近傍領域をマスク領域としてエッジ画像のマスクを行う。この時マスクする領域は、対象とする画像の縦方向画素数 h に対して、 $0.1h$ ピクセル分とする。ここで、連続画像間ではカメラは大きく移動しないと仮定すると、あるフレームにおける各直線はマスク後のエッジ画像に含まれていると考えられる。ここでマスク後のエッジ画像から尤もらしい直線を検出するために、エッジ領域に含まれる画素群を対象としてRANSACと最小二乗法を適用し直線の検出を行う。また、検出された各直線を連続フレーム間で対応づけるために、図2に示す直線検出処理は各直線毎

に行う。

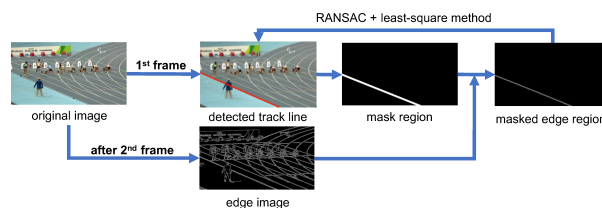


図2 トラック直線検出の流れ

2.1.2 画像間マッチング

続いて検出された直線情報を用いて画像間のマッチングを行う。ここでの目的は各フレームから1フレーム目へ平面射影変換を行うようなホモグラフィ行列を求めることである。提案手法では、まず連続する2画像間でのホモグラフィ行列を求め、その後各ホモグラフィ行列をかけ合わせることで各フレームから1フレーム目へのホモグラフィ行列を求める。しかしながら、連続する2画像間において、検出された直線が重なるような平面射影変換は無数に存在しホモグラフィ行列を一意に定めることはできない。そこで提案手法では連続画像間におけるカメラの移動は平行移動に近似できると仮定することで画像間のトランスレーションを求め、そのトランスレーションを元にホモグラフィ行列を一意に決定する。

まず、連続する2画像間におけるSSD (Sum of Squared Difference) が最小になるような平行移動を求めることでトランスレーションを取得する。この時、陸上ト

トラック平面に対するカメラの移動量を求めるために画像中のテロップ領域, OpenPose[9] による人検出領域は事前にマスクしておく. また SSD を求める際には画像中人物の影が SSD の値に影響を及ぼすことを防ぐため影の影響を受けづらいう, 明度画像を用いる.

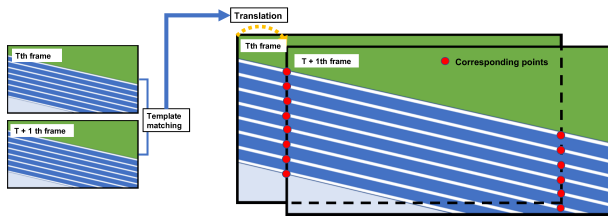


図3 画像間対応点検出の流れ

ここで, ホモグラフィ行列を求めるための対応点の検出方法を説明する. 図3に対応点検出の流れを示す. 画像幅を w , t フレーム目から $t+1$ フレーム目へのトランスレーションを $T(t_x, t_y)$ とし, t フレーム目, i 番目の直線上の点 $p_t(p_x, p_y)$ は $p_y = a_i p_x + b_i$ の式で与えられるとした時, 対応点を図3に示すように t フレーム目の点 $p(t_x, a_i t_x + b_i)$ と $t+1$ フレーム目の点 $p(0, b_i)$ の組, t フレーム目の点 $p(w, a_i w + b_i)$ と $t+1$ フレーム目の点 $p(w - t_x, a_i(w - t_x) + b_i)$ とする. この対応点群を用いて t フレーム目から $t+1$ フレーム目へのホモグラフィ行列 H_t を求める. その後, 各フレームから第1フレーム目へのホモグラフィ行列 H'_t を求める. H'_t は第1フレームまでのホモグラフィ行列の掛け合わせで求まるため以下の式で表すことができる.

$$H'_t = \prod_{t=0}^n H_t \quad (1)$$

その後, 求めたホモグラフィ行列を用いてトラック全体のパノラマ画像を生成する. 図4に生成したパノラマ画像を示す.

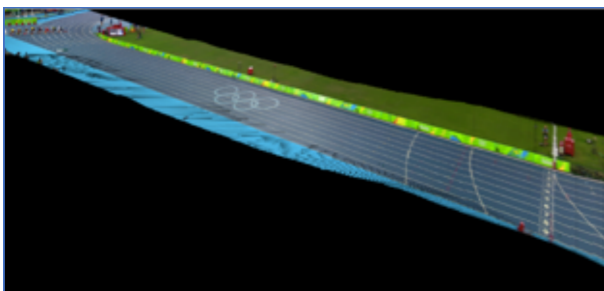


図4 生成したパノラマ画像

2.2 選手画像の投影

続いて選手画像の投影処理について述べる. 本処理は以下に示す4つの処理から成る.

1. 対象選手のトラッキング
2. 映像間でのホモグラフィ行列推定
3. 選手領域画像の投影
4. 画像の合成

まず初めに投影対象選手のトラッキングを行う. 本手法では画像中の人物関節検出モデルである OpenPose[9] を用いて選手の検出・トラッキングを行う. まず OpenPose[9] を用いて各フレーム中の選手を検出する. この時, 第1フレーム目において対象選手周辺領域をクリックすることによって投影対象となる選手を決定する. 2フレーム目以降のトラッキングには OpenPose[9] によって検出された関節位置のうち, 連続する画像間において移動が小さいと考えられる体幹部分(右目, 左目, 鼻, 首)を用いる. 前フレームでのトラッキング対象選手と本フレーム中の選手を比較し, これらの点の距離平均が最小となる人物を本フレームでのトラッキング対象人物とする.

続いて映像間でのホモグラフィ行列 H_p を求める. H_p は第4章で生成したパノラマ画像中に映る100mトラックの角4点を対応点として求める. この時角の4点は, 各映像中のトラック直線と, スタートライン・ゴールライン上を2点クリックすることで求めた直線との交点を求めることで決定する.

続いて事前に求めた H'_t , H_p を用いて選手領域の投影を行う. ここで投影対象選手が映る映像を V_1 , 投影先となる映像を V_2 とした時, V_1 の t フレーム目を V_2 の t フレーム目に投影する様なホモグラフィ行列 $H_{V_1 V_2 t}$ は

$$H_{V_1 V_2 t} = H_{V_1 t'} H_p H_{V_2 t'}^{-1} \quad (2)$$

と表される. ここで $H_{V_1 t'}$ と $H_{V_2 t'}$ は第4章で求めた, 各映像中の t フレームをパノラマ画像に射影するホモグラフィ行列である.

また, $H_{V_1 V_2 t}$ は複数のホモグラフィ行列の掛け合わせによって求められているため, 各ホモグラフィ行列に含まれる誤差が蓄積されている. そのため投影後の画像と背景にズレが生じたりスケールが合わない場合がある. そこで事前に求めたトラック直線が, 投影先の直線と重なる様に投影位置の補正を行う. ここで補正を行うホモグラフィ行列を H_R とした時, V_1 中の t フレーム目の画像 $I_{V_1 t}$ を V_2 中の t フレーム $I_{V_2 t}$ に投影した画像 $I_{V_1 V_2 t}$ は以下の式で表される.

$$I_{V_1 V_2 t} = I_{V_1 t} H_{V_1 V_2 t} H_R \quad (3)$$

最後に $I_{V_1 V_2 t}$ と $I_{V_2 t}$ を合成する. ここで合成を行う領域は対象となる選手のみを含むことが望ましい. そ

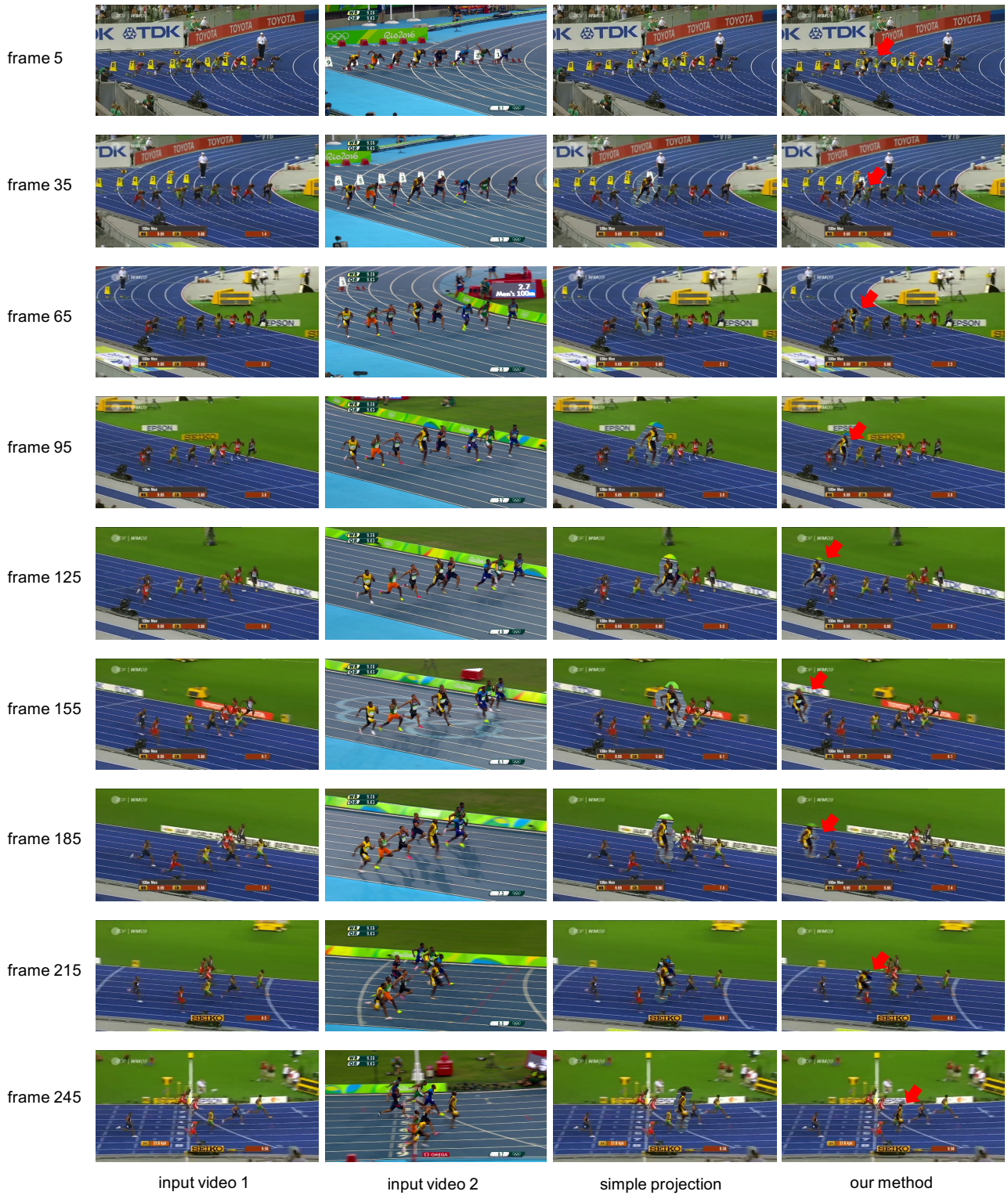


図5 30フレーム毎の入力と仮想対決映像の生成結果

のため OpenPose によって検出された各関節位置を中心とする半径 r の円の和集合を選手領域としマスク画像を生成する. 実験では r の値は実験的に 12px とした. また, より自然な映像を生成するため Poisson Image Blending[10] を用いて境界線が目立たない画像を生成する. poisson image blending[10] を行う際, $I_{V_1V_2t}$ と I_{V_2t} をそのまま入力すると選手領域もボヤけてしまう. そのため poisson image blending[10] の入力, $I_{V_1V_2t}$ 中の選手領域を I_{V_2t} に重畳した画像と I_{V_2t} とする, また境界線をボカすためには, poisson image blending[10] に用いるマスク領域は人領域としたマスク領域よりも大きい物である必要がある. そのため r の値を 10% 大きくした領域をマスク領域とする.

2.3 実験

実験として時空間の異なる 2 つのレース映像に対して提案手法を適用し仮想対決映像を生成した. この度用いた映像は 2009 年世界陸上競技選手権大会の男子 100m 走決勝と, 2016 年世界陸上競技選手権大会の男子 100m 走決勝である. また入力とする 2 つの映像はスタートした瞬間が第一フレームに来る様事前にトリミングを行なった物である. 画像サイズは 640×320 でありフレームレートは 25fps である. 図 5 に入力とした映像, 対象人物を座標変換を行わずに投影した結果 (simple projection), 提案手法による出力を 30 フレーム毎に示す. また, 提案手法の出力結果中の投影された選手を赤い矢印で示す.

出力結果を確認すると投影対象となる選手は対応するレーンに正しく投影されていた. その一方で投影された選手のスケールが映像中の他の選手と異なっていた. これはトラック直線を用いて映像間のスケール合わせをしたためである. 画像内のトラック直線間の距離と選手のスケールはカメラ位置に依存する. そのためカメラ位置が大きく異なる映像間では投影された選手のスケールが他の選手と合わない場合がある. また選手が不自然に加速したり減速したりするシーンが見られた. これはパノラマ画像生成時にトランスレーションが正しく求められていないためだと考えられる. 本手法は第 4 章で述べた様に連続画像間でのテンプレートマッチングによって選手進行方向のカメラ移動量を求めている. しかしながら陸上トラックには基準となるテクスチャが少ないため正しいカメラ移動量が求まらない場合がある. トラック直線の傾き等の情報を用いて, 精度を向上させることが今後の課題である.

2.4 結論

本論文では陸上競技の 100m 走を対象として, 時空間の異なる選手間での仮想対決映像生成手法を提案した. 本手法はカメラ位置が未知であり, かつズーム・パンを行う様なシーンで撮影された映像を対象としてい

る. 選手を別映像中の正しい位置に投影するためには各フレームにおいて選手が何 m 地点を走行しているか把握する必要があるが, 本手法では直線の連続性を制約として用いることで選手の走行位置・投影位置を求めた. 実験では実際に時空間の異なるレースにおける仮想対決映像を生成し, 対応するレーンに正しく選手が投影されることを確認した. その一方で不自然な加速・減速を行うシーンが含まれ, 選手進行方向に対して投影位置に誤差が生じていることが分かった. 今後は直線の連続性に加えて直線の角度を制約として与えることにより精度を向上させることが課題である.

謝辞

本研究は, AIP-PRISM, Japan Science and Technology Agency (JPMJCR18Y2) の支援を受けたものである.

参考文献

- [1] Intel, True View, <https://www.intel.com/content/www/us/en/sports/technology/true-view.html> Accessed 2018/12/12
- [2] Hawk-Eye, Hawk-Eye, <https://www.hawkeyeinnoventions.com/expertise/broadcast-enhancement> Accessed 2018/12/12
- [3] Klein, Georg, and David Murray. "Parallel tracking and mapping for small AR workspaces." International Symposium on Mixed and Augmented Reality, 225-234 (2007).
- [4] Chekhlov, Denis, et al. "Ninja on a plane: Automatic discovery of physical planes for augmented reality using visual slam." International Symposium on Mixed and Augmented Reality, 1-4 (2007)
- [5] 加藤大一郎. "長野オリンピックにおける新しい映像表現技術「仮想対決システム」; 計測と制御 38(4), 281-284 (1999)
- [6] Ng, Pauline C. and Henikoff, Steven: SIFT: Predicting amino acid changes that affect protein function; Nucleic Acids Research, 31(13), 3812-3814 (2003)
- [7] Alcantarilla, Pablo F. and Solutions, T.: SIFT: Fast explicit diffusion for accelerated features in nonlinear scale spaces; IEEE Trans. Patt. Anal. Mach. Intell, 34(7), 1281-1298 (2011)
- [8] Brown, Matthew and Lowe, David G: Automatic panoramic image stitching using invariant features; International Journal of Computer Vision, 74(1), 59-73 (2007)
- [9] Cao Z., Simon T., Wei S. E., Sheikh Y: Realtime multi-person 2d pose estimation using part affinity fields; arXiv preprint arXiv:1611.08050, (2016)
- [10] Pérez, Patrick and Gangnet, Michel and Blake, Andrew, Poisson Image Editing; ACM Trans. Graph, 22(3), 313-318 (2003)