



Visual Computingと そのゲーム応用について

高野晃洋[†]

まえがき

皆様初めまして、ソニー・インタラクティブエンタテインメント(以下SIEと略記)でR&D業務に携わっている、高野晃洋と申します。

本稿では、まず始めに私の入社経緯を大まかにお話し、Visual Computingと括られる技術領域をゲームという分野へ応用するためにどのようなR&D活動がなされているか、就職後の業務・特許出願で取り扱った技術情報(すでに特許広報などで公開されているもの)を交えて、ご紹介させていただければと思います。

私がSIEへ入社する前、大学・大学院では計測・センシングシステムを取り扱う学科コース・専攻に所属しており、具体的な研究テーマはグローブ型デバイスとウェアラブルカメラを用いた指先位置推定、すれ違う歩行者の端末による相互通信・逐次補正を前提とした歩行者位置推定などコンピュータビジョン領域に近いものでした。後者は陽に視覚情報を使うわけではないですが、Visual SLAMを含む自己位置推定システムを前提とするため、領域に近いと言ってもまあ許されるのではないかと…。

就職先を考える際は、これらの研究経験を活かせる可能性が大きく、また生来から慣れ親しんできたゲーム分野で自身の知見をアウトプットすることで、より良いゲーム体験の創出に繋がられればと思い、SIEに入社しました。2018年に入社しまして、5年間Visual Computing領域でのR&D業務に従事しております。

Visual Computingという括り

今回表題に用いましたVisual Computingという用語ですが、これは非常に幅広い領域を内包して用いられることが多い言葉です。コンピュータビジョンはもちろん、コンピュータグラフィックスや画像・動画処理のようなソフトウェア中心の領域、VR/AR、イメージングインタフェースのようなハードウェア中心の領域どちらも内包している、と言って差し支えないでしょう。いくつかのテックカンパニーや大学ではVisual Computingを名に冠したりリサーチグループを作る、あるいは対外的にVisual Computingという括りで先述の領域に属するリサーチテーマを紹介するといった活動が見られます。最大公約数的に表現するワードとしてVisual Computingが使われることから、本当に「視覚」と「コンピュータ」が出てくれば何でもあり、かなりの緩い括りだと認識しています。

その上で、ゲームを中心としたエンタメ領域でVisual Computing技術を展開できるよう取り組んでいるR&DチームがSIEにあり、私はそこに所属

しております。

ゲーム分野がVisual Computingの応用先、というのは想像に難くないかと思えます。例えばよりハイクオリティで軽量の描画処理を達成するにはコンピュータグラフィックスの知見が必要ですし、プレーヤの動向を取り込む(最もシンプルな例としては、プレーヤを撮影し配信画面のワイプ画面にはめ込む等)にはカメラ等のイメージングデバイスを前提に入力した動画像の処理を求められます。VR/ARについては言わずもがなでしょう。

ゲームを取り巻く技術課題とソリューションは多種多様ですが、今回その実例としてご紹介するのは実際に私が発明者に含まれている公開済みの特許出願情報です。特許出願は原則1年半で公開されますので、これらの特許出願については各種特許検索フォームから確認することが可能です。

実例その1

カメラを用いたHMD表示部分の キャリブレーション¹⁾

まずはベタなコンピュータビジョンの例からご紹介します(図1)。HMDの表示部分にはレンズを含む接眼光学系が搭載されていることは、HMDを装着したことのある方であればわかると思いますが、出力したいグラフィックをそのまま素通しで見る時と同様にディスプレイ表示してしまうと、レンズを使用している以上図1のように歪みが現れてしまうことがあります。

こういったレンズ歪みというのは、カ

[†]株式会社ソニー・インタラクティブエンタテインメント インタラクションR&D部門 開発3部3課

"Visual Computing, Applied to Games" by Akihiro Takano (Section #3, Dept. #3, Interaction R&D Division Development, Sony Interactive Entertainment Inc. Tokyo)

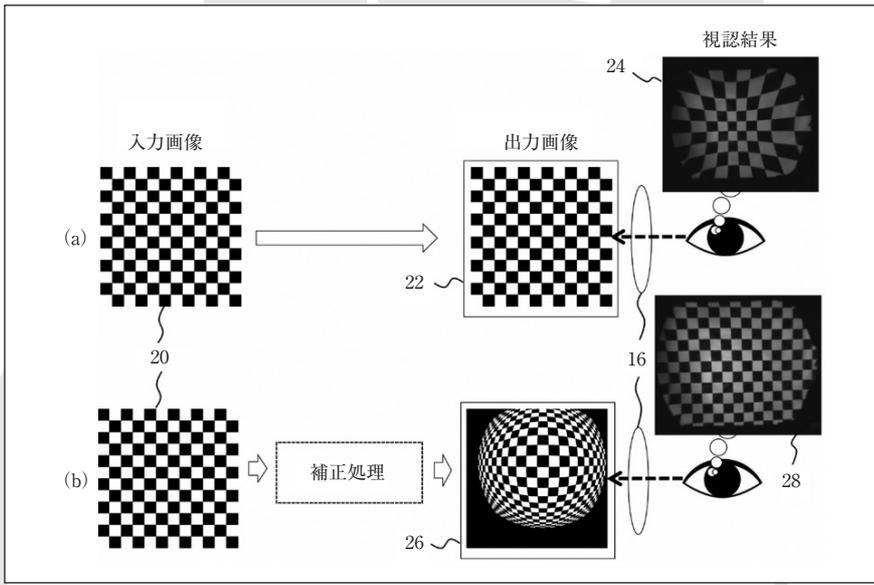


図1 文献¹⁾より引用した特許図面

接眼光学系を通して歪が現れてしまうため、その歪パラメータを元に補正処理をかける必要があります。

メラであればチェッカパターンの平面チャートを複数撮影してカメラパラメータを計算することで除去することが多く、特に有名なのはZhangの手法²⁾でしょう(これは実際に行った経験のある方も多いのではないのでしょうか)。しかしHMDはそうもいかず、仮に同様の方法でレンズをキャリブレーションする、あるいは設計値をダイレクトに使ってしまうとしても、ディスプレイ部分との組付け誤差(光軸位置ズレ・傾きなど)やレンズそのものの個体差から僅かながら歪みが残るケースも考えられます。

そういったケースを踏まえてよりHMDに特化した歪除去のアイデアを具体化したものがこの特許案で、大筋としてはグレイコード、二次元正弦波などのパターン画像出力によってディスプレイ上のピクセル位置と観測用カメラ上の画像座標系の間でマッチングを取り、そのペアを使って歪みを計算するという手法を提案しています。

この特許出願で使われている知見というと、カメラキャリブレーションはもちろんですが、三次元計測で用いられる構造化光もその一つで、パターン画像の表示が三次元物体へのパターン光照射に対応することになります。も

しろん物体計測の場合はプロジェクタのキャリブレーションが済んでいる(パターンの直線構造が保たれている)ことを想定するでしょうから、歪みを計算する用途と完全に合致するわけではないですが、知見の応用は可能です。

実例その2

RGB画像を使った測距情報の補完とフィルタ選択³⁾

ゲーム空間と実世界のインタラクションを視覚情報ベースで行う場合、RGBカメラのみのケースもちろんですが、それに加えて深度をセンシングして奥行方向の情報を使うケースもあります。上で触れている構造化光を用いた3Dセンシング、ステレオカメラによる視差からの深度計算、そしてIR光の反射・到来時間から距離計算をするToF(Time-of-Flight)センサの利用などが考えられますが、IR光を使うケースではRGBセンサの出力と深度情報の解像度がマッチしない場合があります。ステレオカメラ等では物体に遮られた領域の深度が計算できずに欠損領域を生じさせる場合がそれぞれあります。

これらの「RGBデータと比べて足りない深度情報」を埋め合わせる補完手

段も数多く、バイリニア法やバイキュービック法のように深度情報のみから計算する補完手法、重み付きメディアン法やConvolutional Neural Networkを用いてRGBデータから得られる特徴情報を計算に利用する方法などがありますが、本特許出願では手法適用時にConfidence値(深度情報を埋め合わせるのに適した手法を使っているか判定する値)を同時計算することで、より適切な深度マップを作るためにピクセル単位で埋め合わせ手法を選択するというものになります。これによって、例えば、被写体と背景の境界のように本来不連続な部分で連続した深度を持つかのように補間されたり、連続な傾斜面に不連続な段差があるよう補間されたり、といった問題が避けられることを目指したものです。

実例その1・その2のどちらも、ハードウェア側の課題をアルゴリズム側で解決する、という向きのもので、それによって体験価値向上という効果まで繋がる例としてご紹介しました。個々の要素技術まで目線を落としてしまうとゲームと直結するかどうかわかりづらい部分もありますが、そこから一つのシステムに纏め上げてゲーム応用まで繋げていく、というのがわれわれの仕事になります。

実例その3

ゲーム内実況用カメラの最適位置推定⁴⁾

ここまでは実物体を意識した特許出願例でしたが、最後にゲームグラフィックに関する実例の一つをご紹介します(図2)。

昨今eスポーツやゲーム実況・配信などが増える中で、プレーヤとは別にゲーム内で「配信用のカメラマン」が必要となるシーンも増えてきました。

そういったプレーヤ視点と異なる映像を撮るカメラマンの完全自動化、あるいは補助のために、より「盛り上がっている場所」を見つけて移動・提案するというのが本特許出願の肝になります。

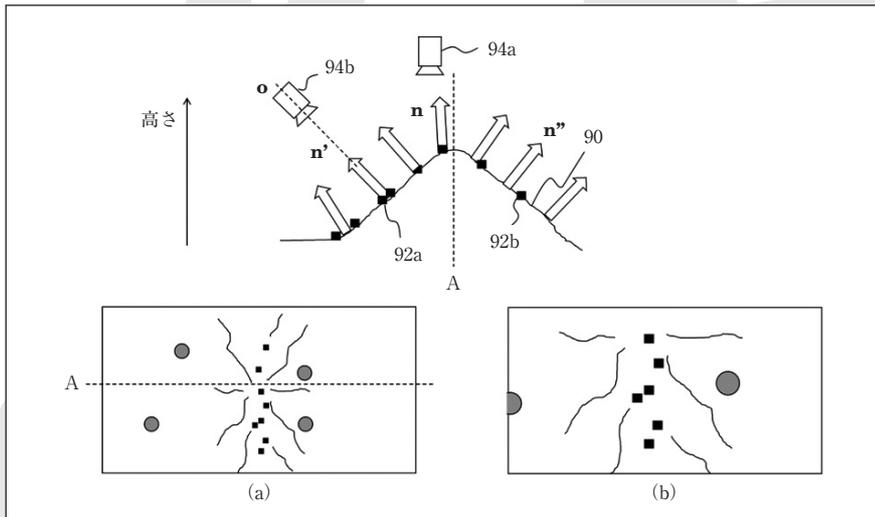


図2 文献³⁾より引用した特許図面

ある斜面にプレーヤ群がいるとして、(a)画面のプレーヤ群がクラスタを形成している場合、クラスタ中心が頂点近く、法線に則ってほぼ直下を映すようにカメラが置かれます。(b)片面のプレーヤ群のみでクラスタを形成している場合、クラスタ中心の法線に則るとカメラ光軸は直下ではなく斜め下方向を向き、その俯角から全プレーヤを収められるポジションへカメラが移動することになります。

この「盛り上がっている場所」というのが非常に定性的な表現ですので補足しますと、例えば、プレーヤ同士の対戦ゲームの試中に多くのキャラクタが密集しているエリアがあるとすれば、そこで試合が動く可能性が高い、だからカメラマンは向かった方がいい、という類推はカメラ位置のプリミティブな提案理由として適切でしょう。そうすると位置は大まかに決められますが、その上で映像の中心をどこに置か、画角をどの程度確保するか、俯角をどの程度に設定するかなどの情報を与える必要もあり、本特許出願に則ると、図2のように全プレーヤのクラスタリング結果から近傍のプレーヤをすべて収められるようクラスタ中心・直径を活用、あとは地形情報から俯角を再計算、という形になります。

この実例は前二つと比べてかなり毛色が異なりますが、これもゲームグラ

フィックの問題ですのでやはり Visual Computing と紐づいた応用実例と言えるでしょう。ゲーム内でどのようなプレーヤ視点を構築すればいいか、という提案はそれこそゲームの数だけなされてきたものですが、プレーヤとは独立した視点をどう決めれば良いかという一つ捻った問題に対してアプローチした特許出願でした。

むすび

以上三つの出願実例から、ゲームを取り巻く技術課題と Visual Computing に基づく解決策、そのそれぞれの幅広さを見ていただけたかと思います。

先述した大学の頃に取り扱った研究テーマと関連が深いのは、実例その1くらいでしょうか。その3に至っては大分かけ離れた問題を取り扱っていましたが、そうは言ってもシステムの裏にある要素技術については、新たに身

に着けた部分より、すでに身につけていた知見(構造化光、数値補間手法、クラスタリング手法など)を応用した部分が大きかったかと、改めて振り返ると思います。特にアルゴリズムについてはさまざまな応用課題に対して通底するな、という実感を持つ5年間でした。とはいえ、昨今の技術的進歩も目覚ましいものがありますので、やはりキャッチアップはおろそかにできませんが。

もし皆様がゲームをされる際、あるいは何らかの題材で取り扱われる際には、その裏でどのような Visual Computing 技術が動いているか、思いを馳せてみるのはいかがでしょうか、馳せるだけでなく実際に調べていただけると、開発者の方々もお喜ばれるかもしれません。何なら「こうすればもっと面白いゲーム体験になるじゃないか」というアイデアをお持ちの学生の方は、その思いをねじ込むつもりでゲーム関連企業の門を叩いてみてはいかがでしょうか。私も有り体に言えばそんな感じでしたので、遠慮なく叩きつけていくことをおすすめします。

(2023年3月1日受付)

【文 献】

- 1) 高野晃洋, 明官佳宏: “キャリブレーション装置, キャリブレーションシステム, および表示装置のキャリブレーション方法”, 特開2020-182127 (Nov. 2020)
- 2) Z. Zhang: "A Flexible New Technique for Camera Calibration", Microsoft Research Technical Report, MSR-TR-98-71 (Dec. 1998)
- 3) 高野晃洋, 明官佳宏, 稲田徹悟: “距離情報生成装置および距離情報生成方法”, WO2022064583 (Mar. 2022)
- 4) 高野晃洋: “実況画像表示支援装置, ゲームシステム, および実況画像表示支援方法”, WO2022130568 (June 2022)