

氏 名	SKOCZEWSKI MATEUSZ
博士の専攻分野の名称	博士（学術）
学位記号番号	博理工甲第 799 号
学位授与年月日	平成 22 年 9 月 17 日
学位授与の条件	学位規則第 4 条第 1 項該当
学位論文題目	Real-time Augmented Reality Based on Neuroevolved Descriptors and Sensor Fusion (ニューロ発展型特徴量とセンサー融合による実時間拡張現実感)
論文審査委員	委員長 教授 久野 義徳 委員 教授 大澤 裕 委員 教授 前川 仁 委員 准教授 内田 淳史

## 論文の内容の要旨

This dissertation presents real time augmented reality system based on local features tracking fused with accelerometer data. Augmented reality is the process of registering projected computer-generated images over a user's view of the physical world. These graphics must be exactly registered to real objects in the scene and this requires AR systems to track a camera in a three dimensional space. With the use of a current mobile smart-phones equipped with camera and gravity sensors, augmented reality can also be experienced on such a devices.

Working in a mobile environment introduces new challenges not previously encountered while generating AR on a standard computers. Those include relatively small computational power and low quality cameras. This dissertation describes a number of novel contributions that improve the state of the art in augmented reality technology and addresses computational complexity issue by providing techniques that do not require on-line severe calculations.

Firstly, we introduce a novel local features descriptor that is based on a neural networks evolution - NEAT (NeuroEvolution of Augmenting Topologies) approach. Our descriptor, coined NELFD (NeuroEvolved Local Feature Descriptor), encodes data around points of interest in the image using a neural network with evolved topology and weights. Due to its adaptive nature and fast computational capability, an evolved neural network has been considered as a new alternative in local descriptor computation. Experimental results are presented for both the standard evaluation set and the real-life acquired images. Our results show that this approach is much faster and comparable in robustness, as well as in some instances, superior to the current state-of-the-art, and thus applicable to the AR systems.

Secondly, we present a novel approach for markerless pose tracking applicable for real-time augmented reality. Visual tracking that consistently locates a desired feature in each image of an input sequence has previously been

applied to AR, however in most of the cases this has used artificial markers placed in the scene. AR approaches based on landmarks are able to perform fast detection and recognition in real-time; however, the detection process is very sensitive to occlusion and requires undesirable fiducial markers to be put where the augmentation takes place. In this paper we apply NELFD as a natural features tracker in order to precisely calculate artificial rendered object position and camera rotation.

We estimate the 3D camera pose by detecting local invariant image features and combining them with the camera's accelerometer data. For every image frame, a correspondence between 2D feature points is calculated and the camera's pose is established based on additional sensor information. Calculating all of the descriptor matches for every frame and comparing them with the referential image would be computationally complex. Thus, we introduce a local features cloud. This local features cloud consists of those descriptors of which the Euclidean distance to the rendering center and strength (Haar response) falls within a specified threshold. Rendering center is decided in advance and represents a place where augmented object will be rendered. The position of the augmented object is determined by position of the cloud. Firstly, object is positioned in the virtual space in a place where the cloud (or due to occlusion, a portion of it) was detected. In the next step we rotate the virtual camera by the angle of the cloud rotation and multiply current projection matrix by the accelerometer matrix calculated from the current gravity vector. We perform a non-uniform scaling along all the object's axes, and the desired scale factor is calculated from the intrinsic camera parameters and the current distance between individual detected local features which belong to the cloud in the image. We used a local features (NELFD descriptors) cloud for matching the current camera frame with the reference frame. Our system was able to work with natural scene features instead of user created artificial landmarks.

In this paper we discuss local image features - interesting points that can be extracted to provide description of the object and computer vision technology used behind AR systems. We shortly introduce vision framework that stands as a base for described approaches. Fundamental methods like corner and edge detectors as well as local grayvalue invariants are presented. SIFT (Scale-Invariant Feature Transform) is described as a robust approach for image local feature generation. We shortly introduced several other descriptors and methods of points of interest localization along with their performance comparison. Augmented Reality technique background is also discussed. We introduce early real-time systems together with visual servoing as well as recent advances in visual tracking. We outline the tracking strategies used in recent AR applications - passive and active fiducial tracking and introduce markerless visual tracking system which will be used as the primary registration sensor (together with additional accelerometer data) for the AR application in this thesis.

Finally, this thesis shows that due to estimation accuracy and low computational complexity our approach can be considered as a new alternative in the real-time augmenting process. Our system robustness has been achieved thanks to the local features cloud that consists of the salient image region descriptors and weights determined by the distance from the object and respective local feature; together with accelerometer readings fusion that gives proper rendered object tilt estimation. Experimental evaluation proved that our method is capable of markerless real-time pose tracking and augmentation in an unconstrained environment.

## 論文の審査結果の要旨

本論文の審査委員会は、平成 22 年 7 月 15 日に論文発表会を開催した。なお、これは平成 21 年 10 月 22 日に開催した同論文発表会の審査内容を受けて開催したもので、前回の審査結果（やや改善の余地あり）の指摘も踏まえて詳細な質疑と審査を行った。その論文発表を含む学位論文の審査の結果、本提出論文を、博士（学術）の学位論文として合格と判定した。以下に審査結果の要約を示す。

近年、現実の世界に有用な情報を付加して提示する 拡張現実感 Augmented Reality (AR) への期待が高まっている。これは、いま目にしている現実の情景に、仮想的な物体映像を重ねあわせることで、医療、生産、さらには広告から娯楽に至る様々な分野で幅広い応用がある。ここで、現実世界への仮想物体の重ねあわせとは、観察者の位置や視点が自由に動く際に、それに応じて、現実世界を映すカメラ画像の所望の位置に、仮想物体を正しい位置・姿勢で描画（レンダリング）することをいい、コンピュータビジョンがその基礎にあって、コンピュータグラフィクスとの融合分野という側面を持っている。

現在、AR システムは様々な形での実装が提案されており、その一部は実現されつつあるが、その多くは、現実世界に対する仮想物体のレンダリングのために、視覚的に分かり易い「マーカー」を貼りつけたり、あるいはゲーム機のように、投影画面近傍に赤外線発信装置を置いてそれを操作ハンドル側で受信することで、画面内の仮想物体の位置姿勢を制御するなど、その装置に特化した動作環境を想定している。本研究は、このような制約を取り除き、特殊な装置を動作環境に想定しない AR システムのモバイル・デバイス上での実現を目指したものである。

本論文は 7 章から構成されている。第 1 章は研究全体の一般的な導入である。

第 2 章では、上述の AR システム一般について、関連研究のサーベイとともに述べ、本研究の寄与の方向、即ち、人工的なマーカーを用いずに、手軽なモバイルデバイスによる撮像を想定した AR システムの実現の有効性・必要性等について説明している。

そのためには、シーンから効率よく画像特徴を抽出して利用する必要があるので、第 3 章では、コンピュータビジョンの関連技術について、最近の研究動向、特に、本研究における提案の基礎となる scale-invariant transform (SIFT; スケール不変変換) 特徴と呼ばれる画像の局所特徴量及びそのいくつかの変形版について述べている。これらは、シーンの見え方の変化によらない局所的な特徴として有用である。しかし、画像の多重分解能表現を基礎にしているため計算量が多く、かつ特徴ベクトルの次元数が大きいので、実時間 AR システムにそのまま適用するのに適さない。

そこで、第 4 章ではその改善版を提案している。本研究の着想は、局所特徴の形成をニューラルネットワークのような発展型アルゴリズムによって得ようというもので、Neuro Evolution of Augmenting Topologies (NEAT) 手法を取り入れた「ニューロ発展型局所特徴記述子 (Neuro Evolved Local Feature Descriptor; NELFD)」を提案している。これによって、SIFT ファミリーに比べてコンパクトな特徴ベクトルによって、それらと同等以上の識別能力が得られることを、標準的に用いられている画像データベースと一般シーンの画像で示している。

第 5 章では、AR システムのための局所画像特徴の利用法について述べている。シーンから得られる局所特徴は膨大な数になるので、仮想物体をレンダリングすべき位置から一定距離以内にある一つの平面上の特徴を組み合わせ feature cloud と呼ぶグループを構成する。最初の画像フレーム（基準フレーム）で feature cloud を決めれば、以下の画像フレームでそれに対応する部分を見出すことができ、動くカメラに応じた姿勢変化も計算できる。また、feature cloud により、多少の隠ぺいによって特徴点がいくつか失われても仮想物体のレンダリング位置は安定して抽出できる。

第6章では、前2章で述べたアルゴリズムの、モバイル・デバイスへの実装を示している。モバイルデバイスとしてはスマートフォンのようなものを想定しているが、本論文では、普及型のUSBカメラを3軸の加速度センサと組み合わせて構成している。カメラから得られる画像は毎秒30フレーム、加速度センサからのデータは100Hzで取り込む。低速・大容量の画像データと高速な加速度データを組み合わせて、毎秒10～20フレームの仮想物体レンダリングを実現している。最後に第7章で以上の成果をまとめ、残された課題について論じている。

以上のように、本論文の内容は、学術的に意義のある研究であると評価できる。よって、当学位論文審査委員会は、本論文を博士（学術）の学位論文として合格と判定した。