

1. 担当 PM

稻見 昌彦（東京大学 先端科学技術研究センター 教授）

2. クリエータ氏名

中野萌士（奈良先端科学技術大学院大学）

堀田大地（電気通信大学）

3. 委託金支払額

2,304,000 円

4. テーマ名

VR 空間における食体験の構築

5. 関連 Web サイト

- ティザーサイト：<https://signs0302.github.io/ukemochi/>
- パッケージ配布：<https://booth.pm/ja/items/1822068>

6. テーマ概要

VR 技術が発展し、小説やアニメ、ゲームなどの空想現実の世界を疑似体験することが可能となってきた。現実空間では難しいことも、空想現実では可能である。本プロジェクトは特に食事について着目する。空想現実の世界では、食事制限などで食べられないものや、誰も食べたことがない「ドラゴンの肉」も食べられるはずだ。しかしながら、空想現実空間では食事をする事自体が非常に難しい。ヘッドマウントディスプレイと呼ばれるメガネ型のディスプレイで空想現実空間を体験しているときは、現実世界にある実際の食べ物が見えない。ヘッドマウントディスプレイに装着するフロントカメラの映像などを通して現実世界を見ながら食事をする事は可能であるが、それにより「空想現実世界に自分が存在している」という実在感が失われてしまう。

本プロジェクトでは、空想現実世界での実在感を保ったまま食事を可能にする要素技術を開発した。空想現実世界での実在感を保ったまま、人間が現実空間の物体を認識する方法として、現実世界のものをそのまま表示するのではなく、

空想世界のイメージに合った VR オブジェクトを作成し、現実世界の物体と置き換えて表示する方法がある。例えば、現実世界の電柱をそのまま表示せずに、大自然の中ならば木の VR オブジェクトに、氷河の世界ならば氷柱の VR オブジェクトに置き換えて表示する。本プロジェクトでは、

- (1) 現実の食べ物領域をカメラで検知する
- (2) 食器型 VR オブジェクトを作成する
- (3) 食器型 VR オブジェクトの中に実際の食べ物画像を重ね合わせる

という 3 つの手順を行うことで、空想世界での実在感を保ったままの食事を実現した。加えて、現実世界の食事を空想現実世界の食事に錯覚させる仕組みの実現を目指した。

7. 採択理由

ACM Future of Computing では、「食とコンピューティング」を次世代の情報学のターゲットとして選定している。本申請はこの食とコンピューティングに関するものであり、VR を体験中に現実には存在しない食を提供することを目指している。食を無理やり VR に持ってくるだけでなく、食体験を VR で演出・拡張するような実装もぜひ目指して欲しい。

8. 開発目標

本プロジェクトは VR 空間で没入感や実在感を保ったまま簡単に食事を行うための支援ソフトウェア (Ukemochi) を開発し、新たな食体験の構築を行うことを目的とした。

9. 進捗概要

本システムはクライアントサーバ方式を用いて、クライアントは Unity、サーバは Python を用いて開発されている。使用される HMD として HTC VIVE、アプリケーションとして SteamVR 上で動作するものを想定している。

クライアントは HMD に付属するフロントカメラから現実世界の映像を取得し、SteamVR で動作するアプリケーション上に重畳する役割を持つ。サーバはクライアントで取得された現実世界の映像を入力され、食事領域と食事座標を出力しクライアントへ返却する役割を持つ。

VIVE のフロントカメラから現実映像を取得しサーバに受け渡すことと、サーバから食事領域画像を取得し VR 空間に重畳表示すること、VR 空間の雰囲気にあった表現を行うため、食事領域の大きさに合わせた 3D モデルを生成し、食事領域と組み合わせた表示を行った (図 1)。



図 1：体験者視点からの映像

(左) VR 空間で花見を行いながらカツ丼を食べている様子

(右) VR 空間にあった雰囲気でするため、
現実映像のりんごを 3D モデルのりんごに置換した様子

食事領域の VR 空間に表示

HTC VIVE のフロントカメラから得られた映像をクライアントが取得する。取得した映像を画像データとして、バイナリデータに変換し、Http Web Request として Python で開発したサーバに送信される。サーバは取得した画像から食事領域を計算し、クライアントは MJPEG Streamer を介して食事領域を領域分割した画像を取得する。クライアントは取得した画像の位置情報を計算し、現実映像に重畳する。最後に Steam VR (Open VR) の VR オーバレイ機能をとって、VR 空間に重畳する。

3D モデルを用いた食事表示

最初に、サーバの物体追従器によって推論されたバウンディングボックスの座標をバイナリデータでクライアントは受け取り、Unity 座標系に変換することで、食事領域の中心座標と縦横の大きさを計算する。次に、食事領域の中心座標に 3D モデルを重畳し、計算した縦横の大きさに応じて 3D モデルの大きさを調節する。最後に、食事領域を非表示にすることで食事領域を 3D モデルに置換することができる。例えば現実映像のりんごを 3D モデルのりんごに置換することができる。また、アニメーションオーバーライドを用いることで相手からもりんごを食べているように見えるため、VR 空間での共食が可能である。

サーバ

クライアントによって取得された映像から食事領域のみを抽出するために、サーバは食事座標を出力する物体検出器と食事領域のみを出力する物体追従器の 2 つで構成した (図 2)。サーバ上での食食物体追従の方法として、本プロジェクトは物体検出器で検出された上下左右の食事座標 4 点 (以下、バウンディングボックス) で囲まれる領域に対して、物体追従器がその領域内部の食事領域

とバウンディングボックスを出力させる。クライアントから送信された最初のフレームを(A)物体追従器の初期化に用いて、その後に送信されるフレームを(B)物体追従器を用いた食事トラッキングに用いる。また、サーバ自体の推論は約24FPSで動作する。

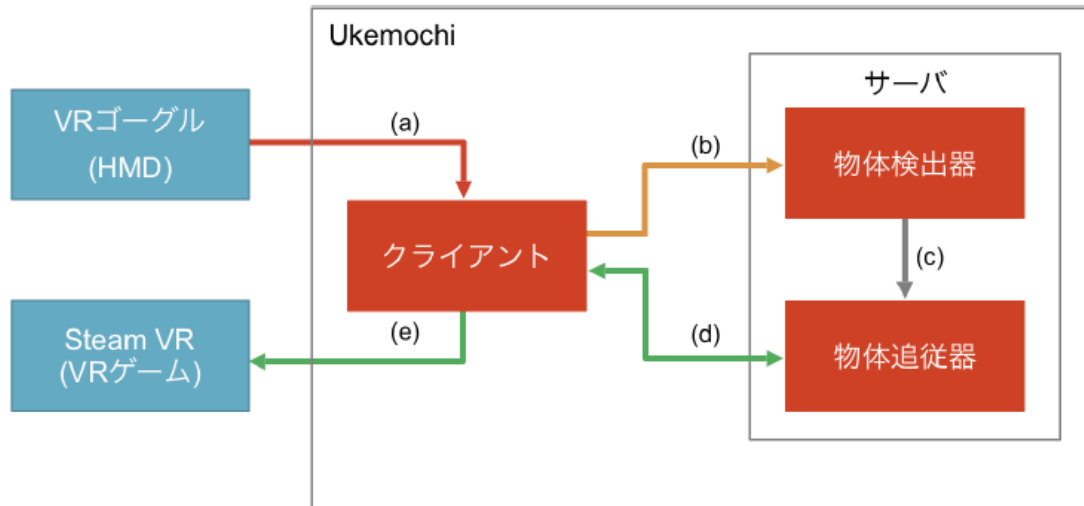


図 2：システム概要

物体検出器

物体検出器の目的は、物体追従器が追従すべき物体を較正するために必要な初期値座標を推論することである。本システムでは、入力画像中の食事を検出するための物体検出器としてYOLOv3を用いた。YOLOv3は解像度320×320の入力に対して、22ms（45FPS）で推論を行うことができる。学習では、UECFood100を使用した。UECFood100は訓練画像9,000枚と評価用画像1,000枚から構築されており、バウンディングボックスアノテーションが付与されている。

物体追従器

物体追従器の目的は、違和感なくVRアプリケーションへ重畳するための食事領域を出力することである。加えて、食事バウンディングボックスも出力することで、クライアントは食事領域に対して3Dモデルを重畳することができる。本システムでは、これらの要件を満たす物体追従器としてSiamMaskを用いた。静止画として入力し領域分割を行うモデルと比較して、SiamMaskは入力されるフレーム中に映る追従物体の相関を用いて領域を出力できるため、より精度の高い領域分割が可能である。加えて、時間軸の特徴量も考慮することができるので、フレーム毎の物体の消失や、推論結果の違いから違和感が生じることを防げる。学習では、VOT2016とVOT2018データセットを使用した。これらのデ

ータセットはテニスの試合といった一般的な動作の動画からの物体追従を行うためのデータセットであるため、食事関連動画は含まれていない。しかしながら、初期値バウンディングボックス内部に含まれる物体を以降のフレームも追従するように学習されているため、本プロジェクトではそのような学習方法を使用し食事物体追従を可能とした。

10. プロジェクト評価

ACM Future of Computing では、「食とコンピューティング」を次世代の情報学のターゲットとして選定している。本プロジェクトはこの食とコンピューティングに関するものであり、VR を体験中に食体験を提供することを目指している。VR とは人間が世界を認識する窓である五感を人工的に再現し提示する技術とも換言できる。現状の VR は五感のうち視覚・聴覚に関しては概ね実現できているが、触覚は研究途上、嗅覚と味覚はほとんど手付かずの状況となっている。本プロジェクトはカメラ付きの HMD を用い、機械学習によって食品のみを切り出し、VR 空間内に投射することに成功した。花見や会食などさまざまなシーンを想定したデモも制作して公開しており、実装力・展開力ともにきわめて優れていると判断する。

11. 今後の課題

本システム『Ukemochi』を紹介した動画は二週間で約六万回再生されており、特に VR コミュニティにおいて注目されている。実装としてまだ粗削りなところはあるが、トラッキングや認識精度を高める方向とは別に、遠隔での共食や食を紹介したコミュニティなど、今後一層の需要が見込める分野において、具体的なサービスにピボットしてゆくことも重要であると考えている。Ukemochi を核としたコミュニティを形成することで、クリエイターら自身の手でなくても広く展開してゆくことが可能となると期待できる。