

Piepple: 拡張現実技術に基づく実世界描画オブジェクトの切り抜き機構とその応用

Piepple: a Mechanism for Clipping a Real-world Drawing Objects based on Augmented Reality and Its Applications

伊藤栄俊[†] 大園忠親[†] 新谷虎松[†]

Eishun Ito, Tadachika Ozono, Toramatsu Shintani

1. はじめに

人間の知的活動における多くの場面で、黒板やホワイトボードなどの無地の平面にアイデアや考えを書き出し、思考を整理することがある。頭の中だけでなく、何かに書き出すことでアイデアや考えを見直すことができたり、気がついていなかった思考に至ることがある。これまでの研究でも、ホワイトボードや付箋に書かれたアイデアを取り出し、発想支援につなげる研究はこれまでに数多く行われている。例えば、Salvatore らの InspirationWall[1] や Shi らの IdeaWall[2] では、ブレインストーミング中における音声を入力情報としたシステムを開発した。Harboe らのシステム [3] では、Affinity diagram の支援システムとして開発された。Harboe らのシステムでは、付箋などの紙に書かれた情報に QR コードを付与し、スマートフォンなどの端末と定点カメラを組み合わせることで情報の読み取りを試みた。こうしたシステムや研究の課題として、アイデア出しという作業中の話題やコンテキストの取得があげられる。我々の過去の研究でも、Mahoca というシステムの中で付箋の抽出を行い、議論のコンテキストとしてアイデアの位置情報の取得するための方法を確立させた [4]。Mahoca における研究の貢献として、AR 技術を用いた現実空間の物体の位置情報の取得があげられる。Mahoca の研究では、付箋中のアイデアという限定した条件での開発であった。本研究ではさらに、AR 技術を用いた位置取得方法を応用し、ホワイトボードに書かれたアイデアを実世界描画オブジェクトとして取り出す切り取り機構を開発した。本稿では、この開発した切り取り機構の機能の説明と、その応用として、オブジェクトの分類と位置情報を考慮した発想支援のための情報推薦モデルの提案を行う。

2. 切り抜き機構

我々は、拡張現実技術に基づく切り抜き機構を用いた実世界型の発想支援システムである Piepple を試作した。本節では現実空間上のアイデアを、実世界描画オブジェクトとして取り出す切り抜き機構の機能の説明を行う。

発想支援を行う際の課題の一つに、発想支援が行われる際の環境の取得が考えられる。こういった状況でシステムによる発想支援が求められたのか判断する必要がある。本研究では、ユーザのアイデアは現実空間のホワイトボードなどの上に記述されている状況を想定している。ホワイトボード上であればペンなどで気軽に書くことができ、テキストだけで

なくグラフィカルな表現で記述することも可能である。そうした情報をシステムが認識する方法として、カメラでの撮影があげられる。カメラでの撮影による取得であれば、安価に用意でき、スマートフォンなどを用いれば、誰でもどこでも利用できるため、アイデア出しを支援することが可能な状況の制約を低く設定することができる。

さらに、撮影によりアイデアを取得するには以下のような課題があげられる。一つ目は、アイデアとそれ以外との分類方法である。画像中にはアイデアのほかにさまざまな情報が含まれている可能性がある。二つ目は、現実空間の位置情報とのマッピングである。アイデアはホワイトボードの様々な場所に記述される可能性がある。また、ホワイトボードのどの位置にどんなアイデアが存在するのかという情報も発想支援では考慮すべきコンテキストである。しかし、ある程度多くのアイデアが描かれた状況でない限り、ホワイトボード上には特徴点が少なく、撮影された画像からだけで位置情報を推定することは難しい。そこで本研究では、拡張現実技術を利用した切り抜き機構を開発した。切り抜き機構の機能は次のとおりである。一つ目は、輪郭に応じた切り抜き機能である。切り抜き方法には 2 種類がある。一つはフリーハンドで輪郭を描き、輪郭に応じて画像を切り抜く方法である。もう一つは、画像処理により物体の輪郭を抽出し、輪郭に沿ってオブジェクトを切り抜く方法である。

画像処理で輪郭の抽出を行う方法について述べる。まず撮影した画像に対して二値化を行う。二値化した画像に対して、輪郭抽出を行う。輪郭情報は点群で表現される。点群に対して、点群全体を含む最小面積の矩形を生成し、一つの切り抜きオブジェクトとして定義する。切り抜きオブジェクトは切り抜き処理前後の画像と、画像平面上での重心を座標としてもち、座標は輪郭抽出の際に得られた座標点群から求める。求めた重心をもとに、現実世界の空間との対応付けを行う。撮影された画像一枚だけでは、撮影されたアイデアの現実空間における 3 次元の位置情報はわからない。そこで、我々が開発した AR 技術を用いた位置情報取得技術を応用する。ARKit¹ という AR 技術を用いたアプリケーション開発を支援するフレームワークがある。ARKit にはマーカレス型の Visual-inertial Odometry という、自己位置認識機能がサポートされている。この ARKit の自己位置認識機能に基づき、撮影画像の 3 次元空間への投影技術を利用して、切り抜きオブジェクトを実世界描画オブジェクトとして現実空間に相当する 3 次元の AR 空間上に投影する。

[†] 名古屋工業大学大学院 工学研究科 情報工学専攻

¹<https://developer.apple.com/jp/arkit/>

3. 切り抜き機構の応用

切り抜き機構により、撮影された画像から切り抜きオブジェクトを取得し、ARKit による現実空間上の位置情報と対応した実世界描画オブジェクトを生成する方法を確立させた。実世界描画オブジェクトは発想支援に利用される。そこで、本節では、切り抜き機構の応用として、発想支援のためのオブジェクトの分類と、位置情報を考慮したアイデア出しを支援する情報推薦モデルに関して述べる。我々は、アイデア出しを支援する情報推薦として、2つの方法を用いた。1つ目は、テキストベースの関連情報推薦である。切り抜き処理を行った際に、オブジェクトの名称を推定しているため、オブジェクトの名称をクエリとした情報推薦を行い、オブジェクトに関連する情報を推薦することでアイデア出しの支援を図った。

3.1 オブジェクトの分類

本研究で取り扱うアイデアはホワイトボード等に記述されたものとする。具体的に、本研究で対象とするアイデアは次の2種類のものとする。一つは、ホワイトボードに書かれるアイデアには、アイデアを言語化したテキストベースの情報、二つ目は高層などを絵やイラストで表現したグラフィカルベースな情報がある。どちらも画像ベースでオブジェクトを抽出する。オブジェクトを認識する際の課題として次のものがあげられる。一つは、一般的な物体認識に比べ、ホワイトボードに書かれるアイデアは特徴量が少なく、判別が難しい。そこで、我々は図1のような分類機構を開発した。開発した分類機構では、テキスト検出ベースとした。オブジェクトに対してテキストであるかどうか判定を行い、テキストが発見されればテキストオブジェクトとして登録し、テキストが発見されなければイラストオブジェクトとして登録される。テキスト検出にはiOSフレームワークのVision¹のテキスト検出を利用した。

テキストであるオブジェクトは、テキストオブジェクトとして処理され、OCRを用いてテキストの内容を認識する。テキストとして判別されなかったものは、イラストオブジェクトとして処理する。テキストオブジェクトとイラストオブジェクトとで区別することでそれぞれの特徴に応じた認識処理を行うことができるため、オブジェクトの認識の向上が期待できる。イラストオブジェクトは学習済みモデルを用いて物体認識を行う。モデルは複数のモデルから状況に応じた選択式を採用した。例えば、手書きイラストであれば、一般物体認識向けのモデルよりも、手書きイラストをもとに学習したものが望ましい。手書きイラストの認識には、Quick,drawの学習データセットを用いた学習済みモデルを利用した。手書きイラスト以外の物体を認識させる際は学習済みのInception-v3モデルを利用した。これらのモデルを用いた際の分類結果をオブジェクトのモデルとした。

抽出されたオブジェクトはそれぞれ、抽出時の画像データ、分類結果ラベル、3次元の位置情報、存在する平面の識別子、生成された時間を要素として持ち、識別子を用いて登録される。定式化すると、抽出されたオブジェクトを

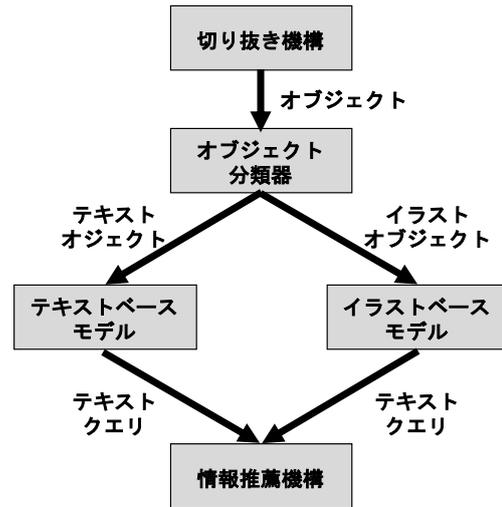


図 1: オブジェクト分類器

$o = \{e_{image}, e_{label}, e_{position}, e_{plane}, e_{time}\}$ と定義する。また、抽出時は必ず平面検出を行うため、検出された平面 $plane$ とユーザ情報 u 、抽出されたオブジェクト集合 O から構成される平面オブジェクト $p = \{plane, u, O\}$ として定義する。

分類されたオブジェクトは、それぞれの特徴に応じた認識処理が行われ、オブジェクトごとにラベルづけが行われる。割り当てられたラベルは発想支援のための情報推薦に利用する。分類結果のラベル情報から情報推薦を行うためのテキストクエリを生成する。テキストクエリには、テキスト情報 $text$ に関連度 r を加えることで情報推薦を行う。定式化すると、 $query = \{text, r\}$ と表す。

3.2 位置情報を考慮した発想支援のための情報推薦

発想を支援するための情報推薦の方法について述べる。本研究では、アイデア出しを支援するための方法として、位置情報を考慮した情報推薦のアルゴリズムを提案する。本研究ではアイデア出しを行なっている最中のコンテキストとして、それぞれのオブジェクトの位置情報に注目した。オブジェクト間の距離を関連性と結びつけて考え、距離や方向に応じた情報推薦を検討した。そこで、ここで我々は重力モデルを提案する。重力モデルに基づく情報推薦では、議論中の平面 p に含まれるオブジェクト全体から議論中の話題に応じた推薦を行う場合と、ユーザがあるオブジェクト δ を選択したオブジェクトに関連する情報を推薦する場合が考えられる。それぞれの状況に関して述べる。

現在行われている議論の話題の推定方法について述べる。現在アイデア出しが行われている時の話題 c とする。我々の提案する重力モデルでは、平面 p 中に含まれるオブジェクトの集合 $O = o_0, o_1, \dots, o_n$ に対して重心を求める。重心は X で表し、式(1)により求める。 N は平面 p 上に存在するオブジェクト数である。 s_i はオブジェクトの重みである。今回は重みとしてオブジェクトの大きさを重みとした。

$$X_p = \frac{\sum_{i=0}^N s_i x_i}{\sum_{i=0}^N x_i} \quad (1)$$

¹<https://developer.apple.com/documentation/vision>

議論中の平面 p における推薦時の話題を c_p とし、次の式で求める。 t_i はオブジェクト o_i の分類ラベルのテキストに対して word2vec を用いて生成したベクトルである。

$$\hat{c}_p = \sum_{i=0}^N t_i (\|x_i - X\|_2) \quad (2)$$

各オブジェクトにはテキストのラベルが付与されているため、word2vec をで分散表現を与えることで、話題に対する関連度をテキストから類似度として求めることができる。加えて、大きさと座標に基づいて重心に相当するものを求めることで、話題に対する影響度を距離によって算出する。話題に対する類似度と影響度の和から、全体の話題をベクトル表現として求める。求めた話題 \hat{c}_p に対して、あらかじめ設定した話題集合 $C = c_0, c_1, \dots, c_m$ から、求めた話題 \hat{c}_p に近い話題を現在の話題として登録する。ゆえに式では次のようになる。

$$c_p = \min_{c_i \in C} \cos(c_i, \hat{c}_p) \quad (3)$$

関連度 r は距離のばらつき度で求める。算出方法は、次のとおりである。

$$r = \frac{1}{\sum_{i=0}^N \|x_i - X\|_2} \quad (4)$$

求めた話題 c_p と関連度 r に対して、テキストクエリを作成し、推薦システムに送ることで状況に応じた情報推薦を実現させる。

選択したオブジェクトに関連する情報推薦に関して述べる。オブジェクトと現在の議論との関連度を r で表現することとする。関連度 r は現在行われている議論の話題 c と、選択オブジェクトの分類結果ラベルに基づくテキストの類似度と、選択オブジェクトの位置情報と話題の重心との距離を用いて決定する。テキストの類似度 $\text{sim}(\hat{o}, o_i)$ には、word2vec で表現された分散表現におけるコサイン類似度を採用する。word2vec は skip-gram モデルに基づく、Wikipedia の全記事を元にした学習済みモデルを利用する。選択オブジェクト \hat{o} に対して、平面状のオブジェクトの部分集合 $O_s = o_0, o_1, \dots, o_n$ に対して、部分集合の重心と関連度を求める。重心 X_s は次の式で求める。

$$X_s = \frac{\sum_{i=0}^n w_i x_i}{\sum_{i=0}^n x_i} \quad (5)$$

位置情報の関連度 r は、推薦時にユーザが使用している平面 p をもとに決定する。

$$r = \frac{1}{\|x_s - X\|_2} \quad (6)$$

x_s は選択オブジェクト \hat{o} の座標で、 X_s は平面上のオブジェクトの部分集合 O_s の重心である。 k は定数で、調整することが可能である。 w_i はオブジェクトがもつ重要度を表す。 w_i はテキスト類似度 $\text{sim}(\hat{o}, o_i)$ とオブジェクトの重み s_i の積で求める。

$$w_i = s_i \text{sim}(\hat{o}, o_i) \quad (7)$$

求めた関連度 r は、選択オブジェクト \hat{o} の分類ラベルのテキストとともに関連情報を選択する際のパラメータとして用いる。

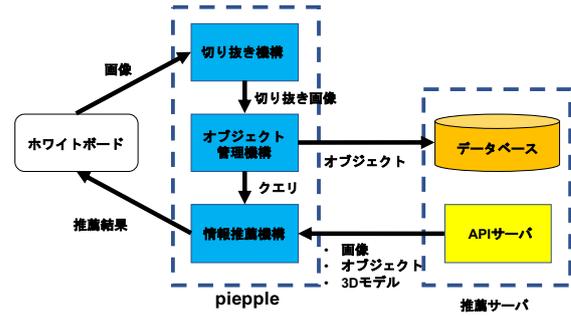


図 2: システム構成

本システムで行う情報推薦では、分類ラベルをもとにした画像検索による関連画像の推薦、同じく分類ラベルをもとにした 3D モデルの推薦、および位置情報を考慮したオブジェクトの推薦の 3 種類である。まず、画像検索に基づく関連画像の推薦について述べる。関連画像の検索では、Microsoft Azure の Image Search API² を利用した。検索結果の画像は透明な板状の 3 次元モデルに投影したものを提示する。3D モデルの推薦では Google の Poly API³ を利用した。Poly API ではクエリを元に 3 次元モデルを検索することができる。3D モデルには様々な形式が存在するが、今回の実装では、ARKit で利用することができる obj 形式のものだけを取得するようにした。過去のオブジェクトの推薦では、推薦サーバに保存されたオブジェクトの情報を元に、類似した関連をもつオブジェクトを推薦する。それぞれの推薦結果から関連度 r に応じてランダムに情報を選択する。関連度は距離が大きいほど小さくなり、 $0 \leq r \leq 1$ を満たすので、検索結果から上位 $r\%$ を関連情報としてランダムに k 個選択する。

4. システム構成

本研究で開発した pipple のシステム構成について説明する。図 2 にシステム構成を示す。本システムは、切り抜き機構、オブジェクト分類器、情報推薦機構からなる。切り抜き機構では、ユーザがホワイトボードに書かれたアイデアを撮影した際の画像を入力として、画像の切り抜き処理を行い、オブジェクトを生成する。生成されたオブジェクトはオブジェクト管理機構に送られ管理される。各オブジェクトは情報推薦および共有用のサーバにあるデータベースとリンクづけられ、保存される。生成されたオブジェクトからクエリを作成し、情報推薦機構に送り発送支援を行うための様々な情報を用意する。送られてくる情報には、画像、3D モデル、および過去のオブジェクトである。

図 3 は、本システムの実行画面を示している。本稿に関連する機能とインターフェースについて説明する。まず、画面右の領域に配置されているボタンはオブジェクトを生成したり、オブジェクトを検索、もしくは情報推薦を促す際に利用する。画面下部の領域に配置されているボタン群はオブジェクトを

²<https://azure.microsoft.com/ja-jp/services/cognitive-services/bing-image-search-api/>

³<https://developers.google.com/poly/develop/api>

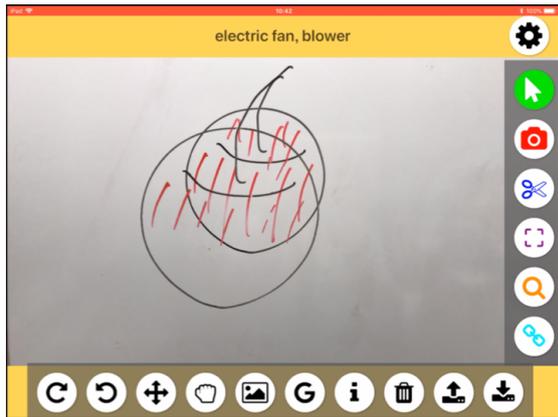


図 3: Piepple のスクリーンショット

操作するとき使用する．オブジェクトを移動させたり，分類ラベルなどの情報を確認する際に利用する．

5. 考察

我々が開発したシステムに関する考察を行う．我々の開発したシステム Piepple の特徴は，ホワイトボード状に描かれたオブジェクトの抽出と，抽出されたオブジェクトの認識とラベリング，そしてラベル情報をもとにした情報推薦である．オブジェクトの抽出には切り抜き機構により実現した．切り抜き機構はタブレット端末のカメラで撮影された画像に対して閾値を定めて自動抽出を可能にした．抽出する対象を黒の輪郭をもつ物体に限定しているため，現在はカラー画像をグレースケールに変換してから二値化を行っている．しかし，様々な色で描かれたものや暗い環境でも対応できるように，状況に応じた閾値の設定を可能にしたい．例えば，赤，青，もしくは黄色などの色で描かれたアイデアを抽出する場合は，HSV の色空間における色相 H で敷地を定め，フィルターなどをかけられる方が望ましい．

抽出したオブジェクトはオブジェクト分類器を用いて分類し，ラベル付けを行い管理している．オブジェクトを分類する際は，テキスト検出を行い，テキストが検出されたものをテキストオブジェクトと分類し，そうでないものをイラストオブジェクトと分類する分類器を採用した．この分類器はテキスト検出の精度が重要となる．現在は Vision とうフレームワークから提供されている API を利用しているが，今後精度の向上を行うためには検出用のモデルの作成が検討される．また，イラストオブジェクトに対しても学習済みモデルを用いて分類を行っているが，こうした学習済みモデルの調整により精度の向上を図っていきたい．

本システムの貢献として，位置情報を考慮した情報推薦のための重力モデルの提案がある．重力モデルでは平面 p 上に存在するオブジェクトの大きさや位置情報をもとに現在の議論の話題 c_p の推定と重心 X の算出を行った．話題は事前に用意したものに適したものを選択する．例えば，同じようなアイデアが固められて描かれていた場合，議論の中心は固めて描かれたアイデアに近いということになる．異なるア

イディアが同じ場所に固められている場合は，その時の議論にはまだ方向性が定まっていないということになる．また，重力モデルでは重心という概念を導入した．重心を定めることでオブジェクトの特徴を位置情報と関連させて用いることができる．オブジェクトの大きさだけでなく，切り抜いたあとの形や，インタラクティブなインターフェースを用いてオブジェクトを編集可能なように設計し，操作された頻度などに応じて定めることで，議論におけるオブジェクトの重要度をテキスト以外の情報として取り入れることが期待される．

しかし，我々の提案するこのモデルにはまだ課題がある．まず解決しなければいけない課題として，本モデルの有効性である．本モデルが有効かどうか評価するためには実際に多くのユーザに利用してもらうことである．だがしかし，そもそも有効なアイデアとは議論ごと個人ごとに異なる．ユーザの傾向や始まりの議論のテーマ等により異なってくるため，多くのユーザが支援されたと感じただけでは有効なモデルが評価できるのかは議論が必要となる．そこで，現在我々は，発想支援のためのデータセットの作成を検討している．

6. おわりに

AR 技術を用いた位置取得方法を応用し，ホワイトボードに書かれたアイデアを実世界描画オブジェクトとして取り出す切り取り機構を開発した．本研究の貢献として，切り抜き機構の応用として，発想支援のためのオブジェクトの分類と，位置情報を考慮したアイデア出しを支援する情報推薦モデルとして重力モデルの提案があげられる．切り抜き機構と情報推薦機構を併せ持つ発想支援のためのシステムとして Piepple は開発された．今後は，Piepple の評価と重力モデルの最適化を行っていく．

謝辞

本研究の一部は JSPS 科研費 15K00422, 16K00420 の助成を受けたものです．

参考文献

- [1] Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. "InspirationWall: Supporting Idea Generation Through Automatic Information Exploration," *Creativity and Cognition (C&C '15)*, pp.103–106, 2015.
- [2] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma, "IdeaWall: Improving Creative Collaboration through Combinatorial Visual Stimuli," *Computer Supported Cooperative Work and Social Computing (CSCW '17)*, pp.594–603, 2017.
- [3] G. Harboe, "Understanding and augmenting a paper arrangement-based method," *Pervasive and ubiquitous computing adjunct publication (UbiComp'13 Adjunct)*, pp.343–348, 2013.
- [4] 伊藤栄俊, 大園忠親, 新谷虎松, "紙付箋抽出における位置取得と解像度を両立する付箋抽出機構の試作," 2018 年度 人工知能学会全国大会 (第 32 回), 2018.