

Wappen:アノテーション共有を用いた複数端末連携システム

前田航平^{†1} 福原知宏^{†1†2} 山田剛一^{†1} 阿倍博信^{†1} 増田英孝^{†1}

概要: 本研究では、外部の情報を参照しながら行う作業の効率化を図るために、複数の端末を用いて連携を行い、情報の参照支援を行うシステムを提案する。その一例として、ユーザに2つの機能を提供するシステムである Wappen を開発した。1つ目はメイン端末上の任意の場面对してアノテーションを付与することができる機能である。2つ目はメイン端末内で任意の場面にいる時、その場面で付与されたアノテーションをサブ端末で参照できる機能である。Wappen は表示画面の類似性を判断して場面を判定する。そのためユーザ毎に別の内容が表示されるようなコンテンツに対して特に有用である。今回は、関連情報を確認しながら行うことが多く、ユーザ毎に別の内容が表示されるコンテンツであるモバイルアプリケーションをターゲットに設定し、本システムの設計と開発を行った。

1. はじめに

何か作業を行う際、その内容に関連した情報を確認しながら行うことは多い。特にスマートフォンのように、画面が小さく複数のモニタを接続できない携帯端末では、関連情報を調べるたびにアプリケーションを切り替える必要があり、元の作業が中断されてしまう。さらには元の作業に戻ったときに、調べて覚えたはずの情報が頭に入っておらず、再度調べ直す場合もある。これらの問題から複数の端末を使用して作業を行うユーザも増えている。複数の端末を用いることで、元の作業を行なっている端末内でのアプリケーション切り替えや、関連情報を記憶する負担を減らすことができる。一方で、ユーザは別の端末で関連情報を調べるという作業が残っており、本来の作業が中断されるという問題は解決されていない。

本研究では、こうした外部の情報を参照しながら行う作業の効率化を図るため、複数の端末を連携させ、情報の参照支援を行うシステムを提案する。その一例として、ユーザに以下の2つの機能を提供するシステムである Wappen を開発した。

- (1) メイン端末上の任意の場面对してアノテーションを付与することができる機能
- (2) メイン端末内で任意の場面にいる時、その場面で付与されたアノテーションをサブ端末で参照できる機能

本論文の構成は以下の通りである。2章では関連研究について述べる。3章では本研究のコンセプトについて述べる。4章では提案システムの設計を述べ、5章で実装について述べる。6章でまとめと今後の展望について述べる。

2. 関連研究

複数の端末を連携させ、情報の提示や作業の補助を行うための様々なインタラクション手法が提案されている。三

田らは端末同士を重ねる操作で端末同士を連携させ、端末同士が協調して情報の可視性を向上させる手法[1]を提案した。栗原らはホバー機能を搭載したモバイル端末のタッチパネルにもう1台の端末を接触させることでホバーイベントが発生し続ける現象に着目し、端末間の連携を容易に行う手法[2]を提案した。

端末連携を利用した様々な協調作業支援に関する研究も行われている。足利らは地図アプリケーションを使用した協調作業に着目し、複数人で地図を操作するときに重要になる事項を調査している[3]。Seifertらは協調作業において、同一画面を複数人で見る方法、個人端末を連携する方法、そして端末とインタラクティブ面で情報をやり取りする方法が問題を解決する手段をどのように変えるか調査した[12]。これらの研究に対し本研究では複数端末を連携させて個人で利用することを想定している。

連携自体の手法ではなく、端末間連携を用いた応用に関しても様々な研究がなされている。工藤らは複数のスマートフォン間を、アプリケーション内のオブジェクトがシームレスに移動できるようにする手法[4]を提案し、試作して

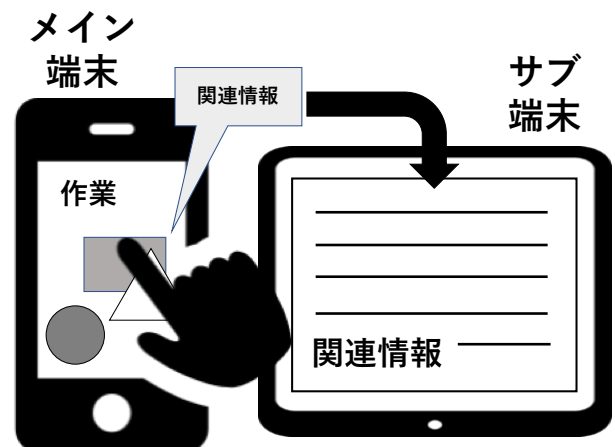


図1 メイン端末とサブ端末

^{†1} 東京電機大学
Tokyo Denki University

^{†2} マルティススープ株式会社
MULTISOUP CO., LTD.

いる。池田らは複数人だけではなく個人のみでの使用を想定した、2 台以上の端末を連携させて、両方の端末を併用するインタフェース[5]を提案している。安本らは複数のモバイル端末を動的かつ立体的に連携させる VISTouch[6]を構築した。Seyed らは複数の端末を一体化させ、必要な機能のみを付属している端末ごと他の人に貸し出せるシステム[13]を実装している。また、杉村らは動画にアノテーションを付与でき、その動画と周囲の家電を連携して、シーンに合わせて照明やエアコンを動作させることにより臨場感を向上させるシステム[7]を開発している。

本研究で開発した Wappen は、モバイル端末上の任意の場面对してアノテーションを付与することができる機能を持つ。これに関連する、Web ページや文書にアノテーションを付与共有する仕組みについては、様々な研究が行われている。高橋らは擬人化会話インタフェースを使った非同期コミュニケーションを行う TelMeA[8]を開発している。佐野らは閲覧者が任意の Web ページ内のコンテンツに対して、付箋によるアノテーションの付与や付箋間へのリンクを貼ることが可能な、Web アノテーションシステム[9]を提案している。Plimmer らは Web ページ上で、複数ユーザーでのデジタルインクアノテーションを提供する iAnnotate[14]を提案している。また、動画にアノテーションを付与する仕組みも研究されている。Guimarães らはビデオに同期コメントを追加できるようにするオンラインビデオコメントシステム[15]の評価を行なっている。Bruna らはリハビリ運動ビデオというユースケースから、モバイルデバイスでキャプチャした動画にアノテーションをつけるツール[16]を開発し評価している。

本研究は端末間連携を用いた応用を行なっているが、モバイル端末自体に手を加える必要のあるものや、複数の端末を同時に操作しなければならないものではない。また、本研究ではアノテーションを付与する機能の開発を行なっているが、Web 上のコンテンツに対してではなく、モバイルアプリケーションにアノテーションを付与するものとなっている点で既存の研究とは違いがある。

3. コンセプト

スマートフォンで行う作業は単一ではなく、ゲームのようなモバイルアプリケーションや Web ページの閲覧など様々である。特にゲームのようなモバイルアプリケーションを行うタスクは、攻略情報などの関連情報を見ながら行うことが多く、本システムを有効に利用できるケースの 1 つである。そのため、Wappen ではメイン端末での作業をモバイルアプリケーションに設定する。アノテーションを付与したいユーザ(付与者)はアプリケーション内の任意の場面でアノテーションを付与でき、アノテーションを参照したいユーザ(閲覧者)はメイン端末に表示されている場面で付与されたアノテーションをサブ端末で参照することがで

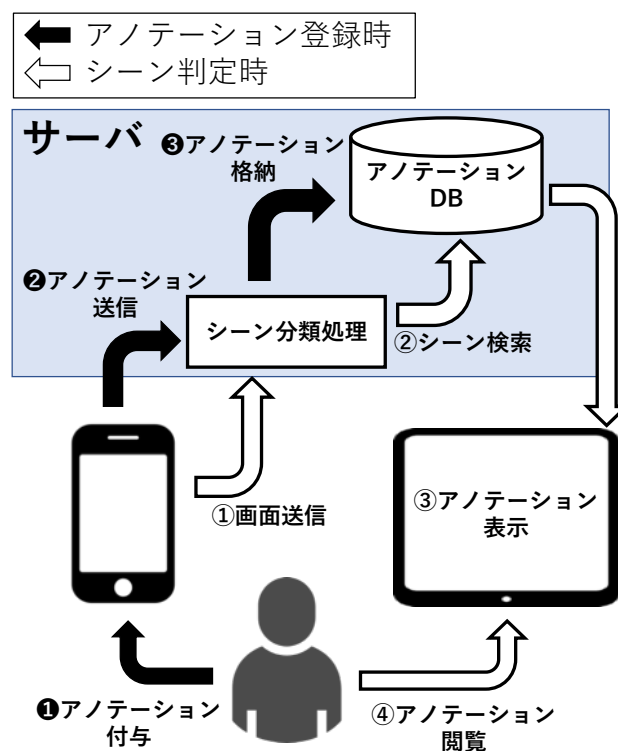


図2 システムの全体図

きる。

本研究では、メインの端末とは別にサブ端末を用意し、メイン端末には行いたい作業、サブ端末には関連情報を表示させる方法を提案する[10][11]。このような複数の携帯端末を用いて連携を行い、情報の参照作業を効率化させる手法としては、以下の2つのアプローチが考えられる。

- (1) メイン端末での表示内容を監視して、閲覧内容に関連した情報をサブ端末に自動表示する方法
 - (2) ユーザに関連情報が必要な場面で、アノテーションという形で関連する情報を付与してもらうという方法
- 前者は、ユーザの操作内容に合わせて自動で定められた関連情報をサブ端末に表示する。この手法は、ユーザにとって必要な関連情報が定まっている場合に有効である。一方、メイン端末の作業内容によっては、サブ端末に表示するべき関連情報の判断が困難になる場合がある。

後者は、ユーザが関連情報を自ら付与することで、自分の欲しい情報を適切な場面で得ることができる。この手法は、サブ端末に表示するべき関連情報がユーザの嗜好や、場面によって変わる場合に有効である。本研究では後者のアノテーション付与によるアプローチを取る。

3.1 複数端末連携のためのアノテーション付与

モバイルアプリケーションにアノテーションを付与するためには、ユーザがアノテーションを付けた時に、モバイルアプリケーション内のどのシーンにいるかを分類して、同じシーンで付けられたアノテーションのみが表示される必要がある。

図 3.1 はシステムの全体像である。ユーザはモバイルア

アプリケーションにアノテーションを付与して、その内容をサーバに送信する。その後、サーバ内でアノテーションはどのシーンで付けられたものかを分類し、その結果とアノテーションの内容を保存する。アノテーションを閲覧したい時には、現在のアプリケーション画面をサーバに送信して、シーンを判定することにより、そのシーンで付けられたアノテーションをサブ端末に表示する。

アノテーションを格納するデータベースは、複数のユーザーで共有できる。そのため、ユーザーは自分で付与したアノテーションだけでなく、他のユーザーが付与したアノテーションも閲覧することができる。これは、特定のグループで協調作業を行う時に有用である。また、個人的な内容をアノテーションとして付与して自分だけで見たい時には、アクセス権限をつけることで対応が可能である。

3.2 ユースケース

今回の複数端末連携では、メイン端末1台に対してサブ端末1台を想定して開発を行った。しかし、本システムのユースケースではサブ端末の台数を制限するべきではない。例えば、アノテーションに関連情報が載っているWebページへのURLが付与されていた場合、サブ端末が1台だと、アノテーションリストとWebページを切り替えるか、両方を同時に表示させる必要がある。同時に表示させた場合では表示領域が狭くなるため、閲覧するときに不便である。そのため、アノテーションリストを表示する端末とWebページを表示する端末を分けることによって、この問題を解決できる場合がある。その場合、表示する関連情報の数によってサブ端末の数を増やす必要がある。

また、メイン端末とサブ端末の役割が入れ替わることもユースケースとして挙げられる。例えば、ユーザーがサブ端末に表示された関連情報に対してアノテーションを付与したいと思った時には、そのサブ端末が今度はメイン端末として動作する必要がある。その際、使用している端末が2台であるならば、メイン端末として使用していた端末をサブ端末として利用することも考えられる。

4. Wappen

モバイルアプリケーションにユーザーが任意のアノテーションを付けて共有し、サブ端末で閲覧することができるツールであるWappenの概要を述べる。

4.1 Wappenの構成

本システムは、メイン端末で動作するWappen アノテータとサブ端末で動作するWappen ウェブに分けられる。Wappen アノテータは、端末のフォアグラウンドに常駐するアプリケーションで、他のアプリケーションが動作している時も使用することができる。図3.1はWappenの概要図である。

Wappen アノテータは2つの機能を持っている。1つ目は、任意のシーンにアノテーションを付与してサーバに送

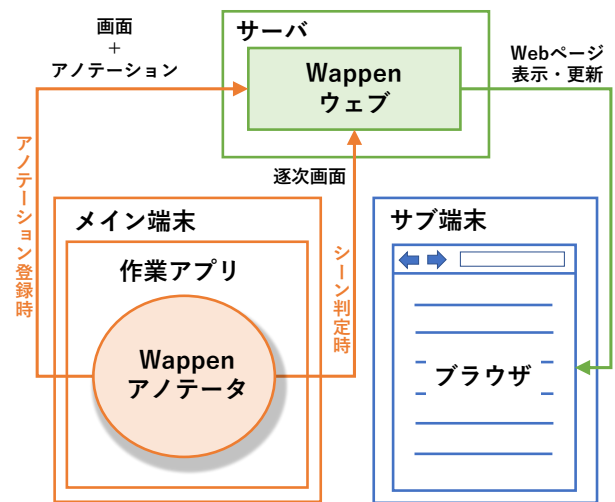


図3 Wappen 概要図

信する機能である。もう1つは画面の逐次送信機能である。この逐次送信は、メイン端末に現在どのシーンが表示されているかを判断するために使用する。またアノテーション付与時と逐次送信時には、どちらも画面のキャプチャ画像をサーバに送信するが、ここで送られた画面はどちらもサーバ内で画像処理によるシーン分類が行われる。

Wappen ウェブでは、メイン端末が現在表示しているシーンと同じシーンで付与されたアノテーションをリスト表示する。

4.2 アノテーションの付与

Wappen アノテータでは、現在表示しているアプリケーションの任意のシーンに対してアノテーションを付与することができる。アノテーションを付与したいシーンで付与ボタンを押すと、アノテーション付与モードになる。付与モードではドラッグで矩形を描くことができ、付与したいシーンのどこに着目してアノテーションをつけるのかを選択する。選択した矩形の内部をタップすることでテキストエリアが表示され、任意のテキストを記入することができる。記入後、付与したテキストと選択した矩形情報、シーンの画面がサーバに送信される。

4.3 アノテーションの分類

Wappen ウェブでは、メイン端末上でユーザーがアノテーションの付与されたシーンを訪れた時に、そのシーンで付与されたアノテーションをまとめて提示する。そのため、ユーザーが付与したアノテーションの中でも、同一シーンで付与されたものをまとめて扱うためにアノテーションをシーンごとに分類する必要がある。

アノテーションが付与されるアプリケーションとWappenは別のアプリケーションのため、アノテーションが付与されるアプリケーションに現在どのシーンが表示されているかをWappen側から取得することはできない。そのためWappenでは、アノテーションが付与されるアプリケーションの画面をキャプチャし、画像処理によって現在表示され

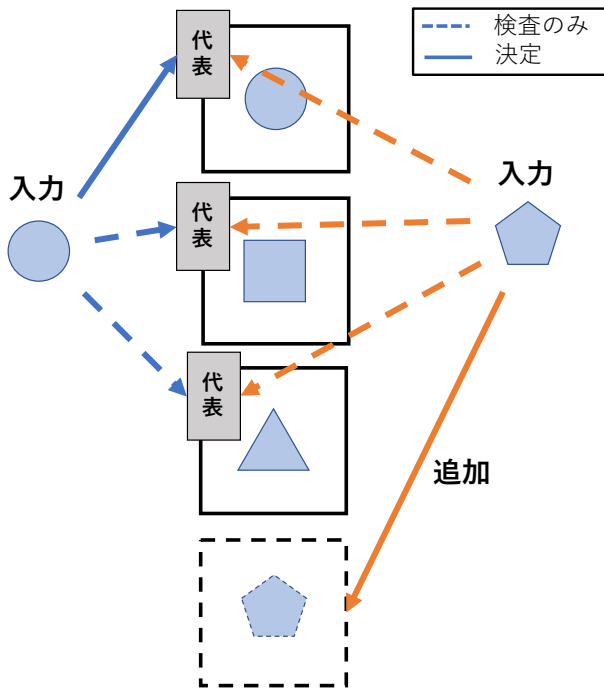


図4 シーン分類のイメージ

ているシーンを特定する。また、アノテーションの分類を行なった後、どのシーンなのかを判別するため、ID を付与しておく(以後シーン ID と呼称する)。これにより、今後分類を行う際、同じシーン ID を持つアノテーションとの比較を 1 回に抑えられる。図 4.2 はシーン分類のイメージ図である。入力図形の種類を 1 つのシーンだと考え、それらがどのシーンと同じかを判定する。入力は各シーン代表と比較して、一番類似しているシーンに分類される。もし既存のシーンに類似しているシーンがなければ、新しいシーンを追加してそこに分類される。アノテーションを分類した時には

4.3.1 シーン振り分け手法

アプリケーションのキャプチャ画像をシーンごとに振り分ける手法について説明する。分類ではキャプチャ画像を入力することで、その画像のシーン ID を出力する。同じシーン ID を持つキャプチャ画像は同じシーンで付与されたキャプチャ画像と判断される。今までに付与された全てのキャプチャ画像と入力画像を照合することは処理速度の関係上困難である。そのため、同じシーン ID を持つ画像の中で代表を決めておき、各シーン ID の代表とのみ、入力画像と照合を行う。これにより照合回数を大きく減らすことができる。また、どのシーン ID の代表とも同じシーンと判定されなかった場合、新しいシーン ID を付与し、新しいシーン代表となる。

分類は特徴点マッチングを用いて行う。まず、入力画像とシーン代表とのマッチング数の偏差値を計算し、極端にマッチング数が少ないものを除く。その後、残ったマッチング数の分散の閾値を使い分類する。

偏差値 T_i と分散 s^2 の計算式は式(1)と式(2)を用いる。偏差値は簡単のため、50 を基準としている。ここで、 s は標準偏差、 n はシーン代表の総数、 x_i はマッチング数、 $\bar{x}$ はマッチング数の平均値となっている。

$$T_i = 10 \times \frac{x_i - \bar{x}}{s} + 50 \quad \dots (1)$$

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \dots (2)$$

特徴点のマッチング数で閾値を決めない理由としては、同一アプリケーションのシーン分類をした際、特徴点のマッチング数に大きな差が出ないため、閾値を決めるのが困難なためである。標準偏差を使用しない理由も、標準偏差は値が小さく閾値を定めるのが困難なためである。

分散が低い場合、入力画像とのマッチング数が特出したシーン代表はいないと判断でき、新しいシーン ID を付与する。一方、入力画像とあるシーン代表が同じシーンの画像だった場合、他のシーン代表とのマッチング数と比較して多いマッチング数を得るため、分散は高くなる。しかし、あるシーン代表が他のシーン代表とのマッチング数と比較して、特出して少ない場合も分散は高くなる。例えば、全く別のアプリケーションのキャプチャ画像がシーン代表として存在する場合である。これに対応するため、分散を計算する前に偏差値を用いて振るいにかける。マッチング数が他と比べて少ないということは、偏差値が他と比べて低いということである。マッチング数が少ないものが、同一シーンではないことは自明である。そのため、偏差値が基準である 50 より低いものはあらかじめ除いておく。除いたことにより、シーン代表が 1 つ残った場合は、そのシーンがマッチング数の特に多いシーン代表であり、入力画像と同一のシーンと判定できる。複数残った場合は分散を計算することで、シーンの分類ができる。

4.4 逐次送信画面のシーン分類

Wappen アノテータには、メイン端末に現在表示されている画面を、逐次的にサーバに送信する機能がある。これは、ユーザが現在メイン端末上でどのシーンにいるのかを判断するためである。逐次画面のシーン分類処理は、アノテーションの分類と同じアルゴリズムで行われる。また、入力された画像のどこに着目してアノテーションがつけられたのかを分類するため、シーン ID が決まった後、同じシーン ID を持つアノテーションの矩形選択画像との比較も必要である。

4.5 アノテーションの表示

Wappen ウェブは、ユニークなユーザ ID で管理されている。これは、Web ページがユーザのメイン端末での操作に連動して動的に変更されるためである。ユーザはサブ端末で Wappen ウェブを開き、Wappen アノテータの逐次送信ボタンを押すだけでサブ端末との連携が完了する。連携をや

めたい時は、逐次送信停止ボタンを押すだけで良い。

メイン端末の逐次送信ボタンが押されると、サーバに現在のメイン端末の画面が逐次的に送信される。サーバはキャプチャ画像を受け取るごとに、シーン振り分け手法と同じアルゴリズムで、受け取った画像のシーン ID を取得する。その後、そのシーン ID と同じ ID が付与されたアノテーションを、Web ページのリストに自動で反映する。

5. Wappen の実装

メイン端末で使用するアプリケーションをモバイルアプリケーションと想定して、アノテーション共有システム Wappen を実装した。システムを実装するにあたり、メイン端末にはスマートフォンを使用した。サブ端末は Web ページが表示できる端末ならば指定はないが、今回はタブレットを用いた。メイン端末の OS には Android を採用している。

5.1 Wappen アノテータの実装

Wappen アノテータはメイン端末で動作するアプリケーションにアノテーションをつけるためのアプリケーションである。Wappen アノテータはフローティングアプリケーションとして実装されている。これは他アプリケーションにいつでもアノテーションをつけることができるという性質上、端末のフォアグラウンドに常駐するアプリケーションである必要性があったためである。

フローティングアプリケーションで実装することにより、他アプリケーションを操作していても Wappen アノテータが表示されるようになったが、他アプリケーションの情報を取得することはできない。そのため、画面のスクリーンキャプチャを行い、サーバ上で画像の解析を行う。スクリーンキャプチャには Android の MediaProjection API を使用した。

アノテーションをサーバに送信する時に送るデータは、ユーザ ID、入力されたテキスト、矩形選択の始点終点座標、画面全体のキャプチャ画像である。現在表示している画面をサーバに逐次送信する時に送るデータは、ユーザ ID と画面全体のキャプチャ画像である。

5.2 Wappen ウェブの実装

Wappen ウェブはサブ端末で表示することでメイン端末との連携が行える Web アプリケーションである。Wappen アノテータから送られてきたアノテーションデータはサーバで前処理が行われ、シーン ID が付与された後、データベースに格納される。今回、シーン ID を付与するために、Python の OpenCV ライブラリを用いて、キャプチャ画像間の SIFT 特徴量[17]を比較することでマッチングを行なっている。そのため、Wappen ウェブは Python の Django を用いて実装した。

サーバで行われる前処理は、送られてきた画像のリサイズや、矩形選択座標を用いた矩形選択範囲のトリミングであ

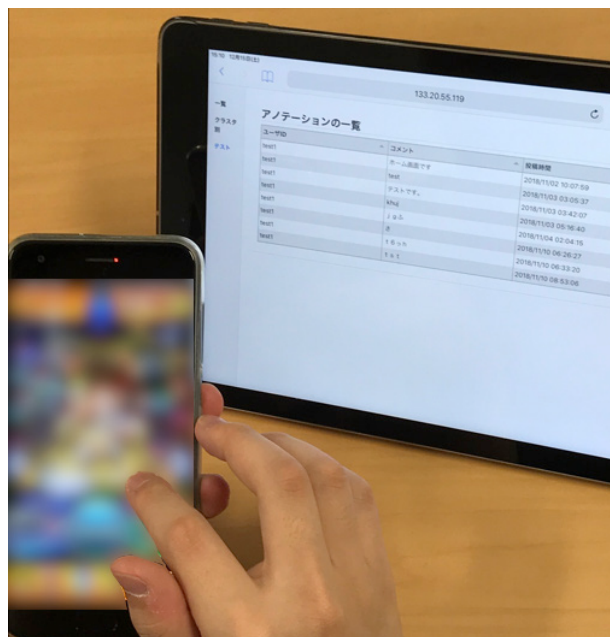


図 5 メイン端末の操作(左)と Wappen ウェブの表示(右)

る。特徴点マッチングの計算量を減らすため、画面全体のキャプチャ画像は 4 分の 1 の大きさに縮小している。シーン ID の付与は特徴点マッチングを用いて行う(4.3.1 項)。

Wappen ウェブでは、1 回の特徴量の比較につき、約 0.20 秒の処理時間がかかる。

Wappen ウェブはユーザ ID で管理されており、ユーザのメイン端末での操作に応じて、表示されるアノテーションのリストが自動更新される。そのためには、Wappen アノテータからキャプチャ画像が逐次送信されてくるたびに、現在メイン端末はどのシーンにいるかを判定して、アノテーションをデータベースから取り出し、その結果を Web ページに自動反映させる必要がある。Web ページへの自動反映には WebSocket を用いた。WebSocket は、Web アプリケーションにおいて、双方向通信を実現するための規格である。サーバとクライアントでコネクションを確立しておくことで、以後サーバとクライアントのどちらからでも通信を行うことが可能となる。これにより、サーバ側でデータベースから取り出したアノテーションをクライアント側にプッシュすることができるため、Web ページに自動で反映できる。図 5.1 は Wappen の使用イメージである。ユーザが操作している端末がメイン端末で、もう一方の端末がサブ端末である。サブ端末には Wappen ウェブが表示されている。Wappen ウェブには、メイン端末に表示されているシーンで付けられたアノテーションがリストで表示されている。

6. おわりに

本研究では、メイン端末とサブ端末を連携させて情報を参照しながら行う作業の効率化を図るため、ユーザに関連情報が必要な場面で、アノテーションという形で関連する

情報を付与してもらい、それを表示するシステムを提案した。その後、メイン端末での作業をモバイルアプリケーションに設定し、モバイルアプリケーションにユーザが任意の場面でアノテーションを付与でき、それを共有できるシステムである **Wappen** を実装した。

モバイルアプリケーションにアノテーションを付与する手法の実現に際し、モバイルアプリケーションの画面を、画像処理を用いてシーンごとに分類する必要があった。

今後は、シーン ID を振り分ける際の特徴点マッチングに使用する特徴量に関しての検証や、処理速度の改善を行った後、評価実験を行い、システムの改良を行う。

参考文献

- [1] 三田裕策, 志築文太郎, 田中二郎: タッチパネル搭載端末同士を重ねる操作による端末連携手法, 研究報告ヒューマンコンピュータインタラクション (HCI), Vol.2013-HCI-154, No.4, pp.1-7(2013).
- [2] 栗原拓郎, 志築文太郎, 田中二郎: ホバー機能に基づくモバイル端末向け連携インタラクション, 研究報告ヒューマンコンピュータインタラクション (HCI), Vol.2014-HCI-160, No.14, pp.1-7(2014).
- [3] 足利えりか, 岩田麻佑, 小牧大治郎ほか: 複数の携帯端末を用いた協調作業における地図操作に関する一考察, マルチメディア, 分散協調とモバイルシンポジウム 2011 論文集 2011, Vol.2011, No.1, pp.587-594(2011).
- [4] 工藤聖広, 辻野友孝, 佐野博之ほか: 複数スマートフォンを用いた分散共有ワークスペースの試作, 第 24 回人工知能学会全国大会論文集, 1D3-3 (2010).
- [5] 池田由優, 赤池英夫, 角田博保: 携帯型情報処理端末を連携したインタフェースの提案と評価, 情報処理学会研究報告ヒューマンコンピュータインタラクション (HCI), 2013-SLP-95, pp.1-2 (2013).
- [6] 安本匡佑, 寺岡丈博: VISTouch – 複数デバイスを用いた動的かつ立体的連携, インタラクション 2015 論文集, C53, pp.926-931(2015).
- [7] 杉村博, 宮澤重明, 数野翔太他: 多種機器連携動画システムとアノテーション共有サービスの開発, 研究報告モバイルコンピューティングとパーベイシブシステム (MBL), Vol.2016-MBL-80, No. 8, pp.1-8(2016).
- [8] 高橋 徹, 濱崎 雅弘, 武田 英明: Avatar-like エージェントを用いた Web コミュニティ支援システム, 人工知能学会全国大会論文集 (第 15 回), pp.1-4(2001).
- [9] 佐野 博之, 浅見 昌平, 大園 忠親ほか: Web エージェントを用いた Web コンテンツへの付箋アノテーションの実現, コンピュータ ソフトウェア, Vol.26, No. 3, pp.69-77(2009).
- [10] 前田航平, 福原知宏, 山田剛一ほか: 主端末作業に関する支援情報表示が可能な補助端末連携システム, 情報処理学会第80回全国大会, 7ZB-01(2018).
- [11] 前田航平, 福原知宏, 山田剛一ほか: 主端末作業に関する支援情報表示が可能な補助端末連携システムの試作, 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2018-HCI-179, No.10, pp.1-8(2018).
- [12] Seifert, J., Simeone, A. L., Schmidt, D., et al.: MobiSurf: Improving Co-located Collaboration through Integrating Mobile Devices and Interactive Surfaces, ITS, pp. 51–60 (2012).
- [13] Seyed, T., Yang, X., Vogel D.: A Modular Smartphone for Lending, Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, UIST '17, pp. 205-215(2017).
- [14] Plimmer, B., Chang, H. S., Doshi, M., et al.: iAnnotate: exploring multi-user ink annotation in web browsers, AUIC '10 Proceedings of the Eleventh Australasian Conference on User Interface, Vol.106, pp.52-60(2010).
- [15] Guimarães, L. R., Cesar, P., Bulterman, C.A. D.: "Let me comment on your video": supporting personalized end-user comments within third-party online videos, WebMedia '12 Proceedings of the 18th Brazilian symposium on Multimedia and the web, pp. 253-260 (2012).
- [16] Bruna C.R. Cunha, Correia, D. R., Maria da Graça Campos Pimentel: Mobile Video Annotations: A Case Study on Supporting Rehabilitation Exercises, WebMedia '15 Proceedings of the 21st Brazilian Symposium on Multimedia and the Web, pp. 245-252, (2015).
- [17] D. G. Lowe: "Distinctive image features from scale-invariant keypoints", Journal of Computer Vision, 60, 2, pp. 91-110 (2004).