

平成22年度

筑波大学情報学群情報科学類

卒業研究論文

題目

実物体への関連付けを利用したデジタルコンテンツ  
参照のためのインタラクション手法

主専攻 ソフトウェアサイエンス主専攻

著者 佐々木慎

指導教員 高橋伸 志築文太郎 三末和男 田中二郎

## 要 旨

我々は生活の中で Web や映像など様々なデジタルコンテンツに関わっている。デジタルコンテンツを直観的に操作するために、実物体にデジタルコンテンツの情報を結び付け、実物体を操作することでデジタルコンテンツを操作する研究が行われている。しかしこれらの研究の多くは特別な物体を用いたり、特別な機器が必要であるなどといった問題点がある。

本研究では、カメラ映像を利用することにより、実物体に関連付けられたデジタルコンテンツを参照することができる手法を提案し、システムの実装を行った。PTZ カメラとジェスチャを用いることにより室内のどこからでも実物体を認識させることができ、その実物体に関連付けられたデジタルコンテンツを参照することが可能となった。またユーザが自由に実物体とデジタルコンテンツの関連付けをすることができるためユーザごとに適した利用が可能となる。

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	ユビキタスコンピューティング	1
1.2	デジタルコンテンツの利用と問題点	1
1.3	本研究の目的	1
1.4	本論文の構成	2
<b>第2章</b>	<b>実物体を用いたデジタルコンテンツ参照手法の提案</b>	<b>3</b>
2.1	アプローチ	3
2.1.1	実物体の利用	3
2.1.2	室内のどこからでも利用できるシステム	3
2.1.3	実物体への自由な関連付け	4
2.2	提案手法	4
<b>第3章</b>	<b>提案システム</b>	<b>5</b>
3.1	システム概要	5
3.2	ジェスチャ	6
3.3	実物体にデジタルコンテンツを関連付ける操作	6
3.3.1	実物体の登録操作	6
3.3.2	実物体へのURLや既存のファイルの関連付け操作	7
3.3.3	実物体への動画像、音声の関連付け操作	8
3.4	実物体に関連付けられたデジタルコンテンツの参照操作	9
3.5	利用シーン	9
<b>第4章</b>	<b>システム構成と実装</b>	<b>11</b>
4.1	システム構成	11
4.1.1	開発環境	12
4.1.2	ネットワークカメラ	12
4.2	処理の流れ	14
4.3	手振り動作の認識	14
4.3.1	広角カメラによる周期的な動作の認識	16
4.3.2	PTZカメラによる周期的な動作の認識	17
4.4	カメラの操作処理	18

4.4.1	広角カメラと PTZ カメラの対応処理 . . . . .	18
4.4.2	ズーム処理 . . . . .	18
4.5	実物体の認識処理 . . . . .	18
4.5.1	実物体の登録 . . . . .	19
4.5.2	実物体の照合 . . . . .	20
4.6	実物体とデジタルコンテンツの関連付け処理 . . . . .	20
4.6.1	実物体と関連付けられたデジタルコンテンツの保存 . . . . .	20
4.6.2	デジタルコンテンツの取得 . . . . .	21
4.6.3	実物体と関連付けられたデジタルコンテンツの参照 . . . . .	21
<b>第 5 章</b>	<b>議論</b>	<b>22</b>
5.1	試用とそこから得られた知見・課題 . . . . .	22
5.2	発展 . . . . .	23
5.2.1	複数地点での使用 . . . . .	23
5.2.2	複数の関連付けやデジタルコンテンツの操作 . . . . .	23
<b>第 6 章</b>	<b>関連研究</b>	<b>24</b>
6.1	身の回りにある実物体を利用した研究 . . . . .	24
6.2	実物体に情報を保存する研究 . . . . .	24
6.3	周期的な動作の認識を活用した研究 . . . . .	25
<b>第 7 章</b>	<b>結論</b>	<b>26</b>
	謝辞	27
	参考文献	28

# 目次

3.1	利用イメージ	5
3.2	実物体の登録操作の流れ	7
3.3	アイコンとメニュー	7
3.4	URL 登録のダイアログ	8
3.5	FILE 登録のダイアログ	8
3.6	デジタルコンテンツの参照操作の流れ	9
4.1	ハードウェア構成のイメージ	11
4.2	構築したシステムの一部	12
4.3	ネットワークカメラ	13
4.4	処理の流れ	15
4.5	広角カメラ映像と低解像度化した濃淡画像	16
4.6	広角カメラでの手振り動作の検出例	16
4.7	PTZ カメラ映像と低解像度化した濃淡画像	17
4.8	PTZ カメラでの手振り動作の検出例	17
4.9	SURF による特徴抽出の例	19
4.10	データベースに保存された実物体のデータ例	19
4.11	実物体の照合の処理の流れ	20
4.12	データベースに保存されたコンテンツのデータ例	21
5.1	試用時に用いたデータ	22
5.2	試用している様子	23

# 第1章 序論

## 1.1 ユビキタスコンピューティング

ユビキタスコンピューティングは、コンピューティング環境の将来像として、Xerox PARC の Mark Weiser によって 1991 年に提唱された概念 [1] である。ユビキタスコンピューティング環境では、日常生活の様々な場所にコンピュータが遍在し、相互にネットワークでつながることで人々の生活を支援する。そしてユーザはコンピュータの存在を過度に意識することなく、その機能を利用することができる。

ユビキタスコンピューティング環境では、誰もがいつでもどこでもコンピュータを利用できるため、従来の、コンピュータの近くに行き操作するというあり方とは異なる。そのためマウスやキーボードといった従来の操作手法で操作するのではなく、新しいインタラクション手法が求められている。

## 1.2 デジタルコンテンツの利用と問題点

我々は日常生活の中で Web や音楽などのデジタルコンテンツと深く関わって生活している。多くの種類のデジタルコンテンツを利用しているが、目的のコンテンツ以外にも多くのコンテンツが存在するため、目的のデジタルコンテンツを検索するのに時間がかかるといった問題点があげられる。

また特定のデジタルコンテンツは利用したい場面が決まっていると考えられる。例えば、朝には今日一日の天気予報や渋滞情報を調べたいといったことや、夜には今日起こったニュースが知りたいなどといったことが挙げられる。時間や状況に応じて利用したいデジタルコンテンツが変わると考えられる。

ユーザがデジタルコンテンツを利用する際には、コンピュータの前に移動し、目的のデジタルコンテンツを検索することが一般的である。また PDA や携帯電話、スマートフォンなどの普及により、小型の端末を用いたデジタルコンテンツの利用も増えてきている。しかしデジタルコンテンツを利用するための端末が近くになければいけないといった問題点がある。

## 1.3 本研究の目的

本研究では、ユーザの身の回りにある実物体に着目し、その実物体にデジタルコンテンツを関連付け、参照することができる手法を提案しそれを実現するシステムを開発することを

目的とする．

## 1.4 本論文の構成

本論文は本章を含め 7 章構成である．以下に各章の概要を示す．第 2 章では本研究のアプローチと提案手法について述べる．第 3 章では提案手法を実現するシステムについて述べ，第 4 章で実際に開発したシステムについて述べる．第 5 章では議論を行い，第 6 章で関連研究について述べ，第 7 章で結論を述べる．

## 第2章 実物体を用いたデジタルコンテンツ参照手法の提案

### 2.1 アプローチ

1章で挙げたデジタルコンテンツの利用や問題点をふまえ、以下に本研究でのアプローチを示す。

#### 2.1.1 実物体の利用

身の回りには目覚まし時計や財布など多くの実物体が存在している。ほとんどの実物体にはそれぞれの使用目的があり使用する場面や状況が決まっている。例えば、朝起きる時に目覚まし時計を使用し、身支度するために歯ブラシ、歯磨き粉を使用するなどである。1章で述べたようにデジタルコンテンツも利用したい状況、時間があると考えられる。そのためデジタルコンテンツを利用したい場面で身近にある実物体を用いることで、そのデジタルコンテンツへのアクセスがしやすくなると考えられる。

実物体を用いてデジタルコンテンツを利用する場合、実物体の本来の使用を妨げず、思った時に即座に利用できることが望ましい。そのためには実物体そのものに特別なセンサや機器などを取り付けずに利用できる方がよいと考えられる。

また実物体を本来の使用目的で利用している場合に、誤ってデジタルコンテンツを利用してしまうといったことを避けなければならない。

#### 2.1.2 室内のどこからでも利用できるシステム

デジタルコンテンツを利用するために特定の場所に移動するよりも、室内のどこからでも利用できる方がよい。そのためには、デジタルコンテンツを利用する意図を明示的に示すこととユーザの位置を把握することが必要になると考えられる。

また室内において普段の生活を行っている場合、ユーザが特別な機器を常に装着しているといったことや、デジタルコンテンツを利用する端末がいつも近くにあるといったことは想定しにくい。

### 2.1.3 実物体への自由な関連付け

ユーザが利用するデジタルコンテンツは様々であり，ユーザごと，状況ごとに利用したいデジタルコンテンツは異なると考えられる．そのためユーザが自由にデジタルコンテンツを関連付けできる方がよい．ユーザが自由に関連付けできることによってそのユーザの生活にそった利用が可能であると考えられる．

## 2.2 提案手法

前節のアプローチをふまえ，デジタルコンテンツの参照手法を提案する．本提案手法ではデジタルコンテンツを身の回りにある実物体に関連付けるということを利用し，その実物体を通してデジタルコンテンツを参照する．具体的には室内に設置したカメラに向かってジェスチャを行い，カメラを通して実物体を認識させることでその実物体に関連付けられたデジタルコンテンツを参照できるというものである．この手法を用いることにより，目的のデジタルコンテンツを室内のどこにいても参照することができる．

## 第3章 提案システム

### 3.1 システム概要

2章で述べた提案手法をふまえた本システムの利用イメージを図3.1に示す。ユーザが室内のどの位置でジェスチャを行っても認識できるようにするために広角カメラを利用し、実物体の認識のためにPTZカメラを利用する。このことによって室内でのどこからでも実物体を認識させることが可能となる。また本システムではユーザ自身が任意のデジタルコンテンツを実物体に関連付けることができ、その実物体を通してデジタルコンテンツを参照することが可能となる。本章の以下の節で具体的な操作の方法を述べる。

なお本システムで参照できるデジタルコンテンツの種類は、動画、音、Webページ、画像などである。

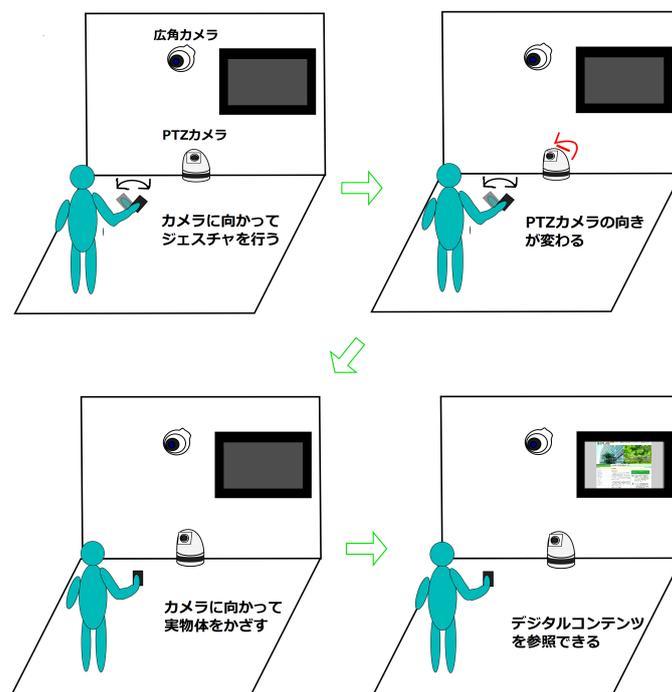


図 3.1: 利用イメージ

## 3.2 ジェスチャ

本研究では、ジェスチャとして、手振りの動作を用いる。なお本研究で用いる手振りの動作とは、手を左右に周期的に動かす動作とする。

我々の日常の生活において手振りは相手と別れる際のあいさつや相手の注意を引く時などに用いられており、身近なジェスチャの一つとして考えられる。また手振りの動作を室内で行うことはあまりなく、室内でのユーザの動作意思を明示的に示すことができると考えられる。

手振りの動作を用いることで、ユーザの位置を特定するとともに、手振りの動作を行っている範囲を PTZ カメラでズームすることにより、実物体の認識も可能となる。

実物体にデジタルコンテンツを関連付ける操作と実物体に関連付けられたデジタルコンテンツを参照する操作の両方で手振りをを用いる。それぞれの操作は手振りをする長さを変えることによって区別する。ジェスチャの種類を 1 つにすることで、ユーザに負担をかけることなく使用することができると考えられる。

## 3.3 実物体にデジタルコンテンツを関連付ける操作

本研究における実物体にデジタルコンテンツを関連付ける操作は、実物体への URL や既存のファイルの関連付け操作、実物体への動画・音声の関連付け操作の 2 つである。

実物体への URL や既存のファイルの関連付け操作では、Web ページの URL やコンピュータ内のファイルのパスを実物体に関連付けることができる。Web ページはユーザがよく利用するデジタルコンテンツとして挙げられる。またユーザが記録したデジタルコンテンツやお気に入りのコンテンツはコンピュータ内に保存するといったことが一般的である。そのため URL や既存のファイルを実物体に関連付けることが必要であると考えられる。

実物体への動画・音声の関連付け操作では、室内に設置したカメラの映像とマイクにより動画・音声を記録し、その内容を実物体に関連付けることができる。動画・音声はメモとしての利用が考えられる。紙やペンなどといったメモを取るための道具がない場合でも、身近にある実物体に動画・音声としてメモを残すことが可能である。

実物体への URL や既存のファイルの関連付け操作と実物体への動画・音声の関連付け操作での実物体を登録する動作は同じ動作を用いる。

### 3.3.1 実物体の登録操作

実物体の登録操作において、まず開始ジェスチャとして手振りの動作を行う。ある一定時間手振りの動作を続けると音が鳴るがその後も手振りの動作を続ける。二回目の音が鳴るまで手振りの動作を続け、音が鳴ったら手振りを止め実物体を PTZ カメラにかざす。PTZ カメラに映った画像が実物体として登録される。実物体が登録されると登録音が鳴り、実物体が登録されたことを確認することができる。また登録時に撮影された画像をディスプレイに表示させることで、ユーザにフィードバックを与えることが可能である。実物体の登録操作の流れを図 3.2 に示す。

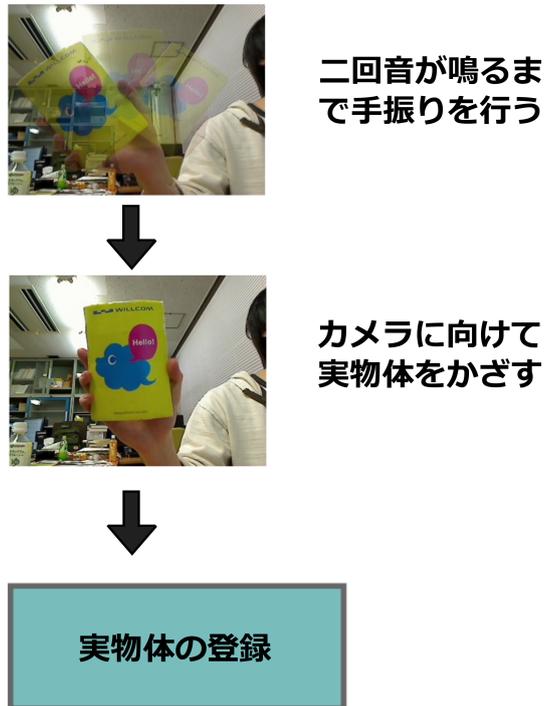


図 3.2: 実物体の登録操作の流れ

### 3.3.2 実物体への URL や既存のファイルの関連付け操作

コンピュータのスクリーン上で目的の Web ページや既存のファイルを選択した後、実物体の登録操作を行うことでデジタルコンテンツと実物体を関連付ける。以下にその詳しい方法を示す。

URL や音楽ファイルなどのコンピュータ内の既存のファイルを実物体に保存するには、まず図 3.3 に示されているシステムトレイのアイコンをクリックする。クリックするとメニューが表示されるので URL を保存する場合は「URL」を、既存のファイルを保存する場合は「FILE」をクリックする。



図 3.3: アイコンとメニュー

URL を選択すると図 3.4 に示されているようなダイアログが表示される。フォームに入力された URL と実物体を関連付けて保存する。現在見ている Web ページと関連付けて保存したい場合は、Get URL のボタンを押すことで、現在 Internet Explorer に表示している Web ページの URL がフォームに入力される。Save ボタンを押すことで、フォームの内容が記録される。その後実物体の登録操作を行うことで実物体と Web ページを関連付けて保存できる。

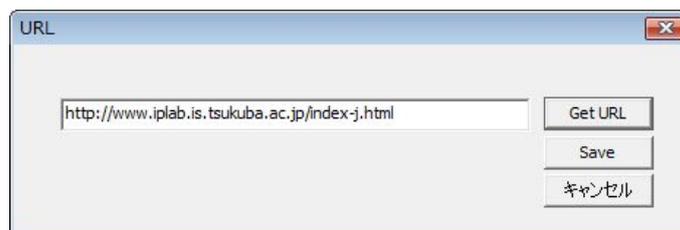


図 3.4: URL 登録のダイアログ

FILE をクリックすると図 3.5 に示されているようなダイアログが表示される。Get Path のボタンをクリックすると、コモンダイアログが現れるのでオブジェクトに関連付けたいファイルを選択する。フォームにそのファイルへのパスが表示されるので、Save ボタンを押すことでフォームの内容が記録される。その後実物体の登録操作を行うことで実物体とファイルを関連付けて保存できる。

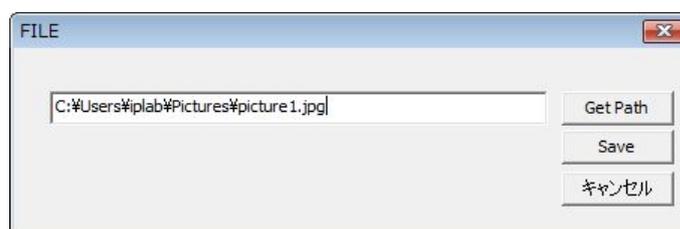


図 3.5: FILE 登録のダイアログ

### 3.3.3 実物体への動画像，音声の関連付け操作

ユーザが実物体に動画像・音声を関連付けて保存する場合は、システムトレイのアイコンをクリックせずに実物体を登録する動作を行う。実物体の登録操作が完了すると動画像・音声の記録が始まる。動画像は PTZ カメラの映像が録画されるので、記録したい音声や身振りなどを行う。手振りの動作を行うと動画像・音声の記録が終了する。また一定時間周期的な動作が認識されない場合は自動的に記録が終了する。

記録された動画像・音声は実物体と関連付けて保存される。

### 3.4 実物体に関連付けられたデジタルコンテンツの参照操作

実物体に関連付けられたデジタルコンテンツを参照するためには、まず開始ジェスチャとして手振りの動作を行う。音が鳴るまで手振りの動作を行い、音が鳴ったら手振りを止め、実物体を PTZ カメラに向かってかざす。かざした実物体が認識させるとその実物体に関連付けられていたデジタルコンテンツが参照できる。関連付けられているデジタルコンテンツが URL ならばその URL の Web ページがブラウザで開かれ、動画像ならその動画像が表示される。デジタルコンテンツの参照操作の流れを図 3.6 に示す。

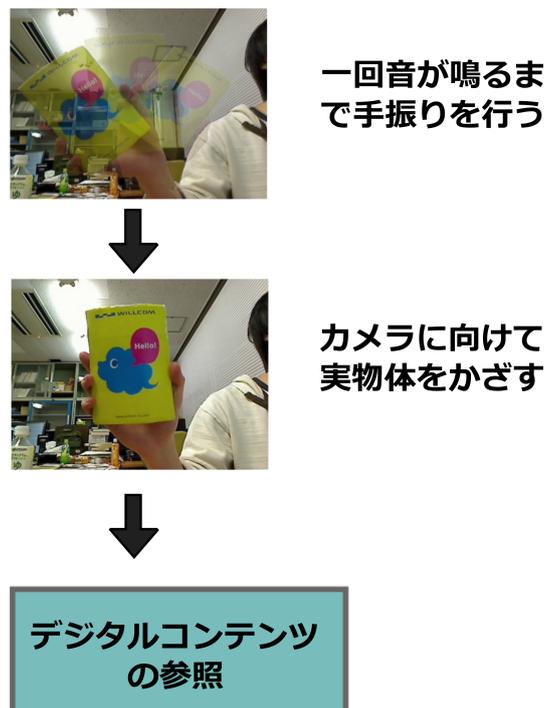


図 3.6: デジタルコンテンツの参照操作の流れ

### 3.5 利用シーン

以下に大学生 A さんの生活をもとに、本システムの利用シーンを示す。  
本システムを用いた A さんのある一日

朝、A さんは目を覚まし大学に行く準備を始めた。まず顔を洗い、歯を磨き始めた。歯を磨いている最中に歯磨き粉のチューブを手を持ちカメラに向かって手振りを行うとディスプレイに天気予報が表示された。その天気予報で A さんはその日の天気を確認し、夕方から雨が降るという予報だったので、傘を持って大学に出かけることができた。

大学から帰宅し部屋で大学の課題を行っているときふと旅行で買った猿の置物が目に入りその時の旅行を懐かしく思った。そこで猿の置物を手に取りカメラに向けて手振りを行うとその旅行中に撮ったビデオが流れその時の旅行をより詳しく思い出すことができた。夜も遅くなったのでベットに移動し眠ろうとした。すると思いがけず研究のアイデアを思いついた。紙やペンが近くになかったため、近くにあった目覚まし時計を手に取りカメラに向かって手振りを行い、思いついたアイデアを動画像・音声として記録した。アイデアをしっかりと記録しほっとした Aさんは安心して眠りについた。

この利用シーンでは

- 歯磨き粉のチューブと天気予報の Web ページ
- 猿の置物とビデオのファイル
- 目覚まし時計と記録した動画像・音声

の3つの関連付けを用いている。

## 第4章 システム構成と実装

### 4.1 システム構成

本システムのハードウェア構成のイメージ図を図4.1に、実際に構築したシステムの一部を図4.2に示す。室内全体での手振りの動作を検出するために広角カメラを、物体認識を行うためにPTZカメラを用いる。そしてデジタルコンテンツを表示、登録するためのディスプレイ、フィードバックを与えるためのスピーカーを壁面近くに設置する。また音声を録音するマイクを設置する。

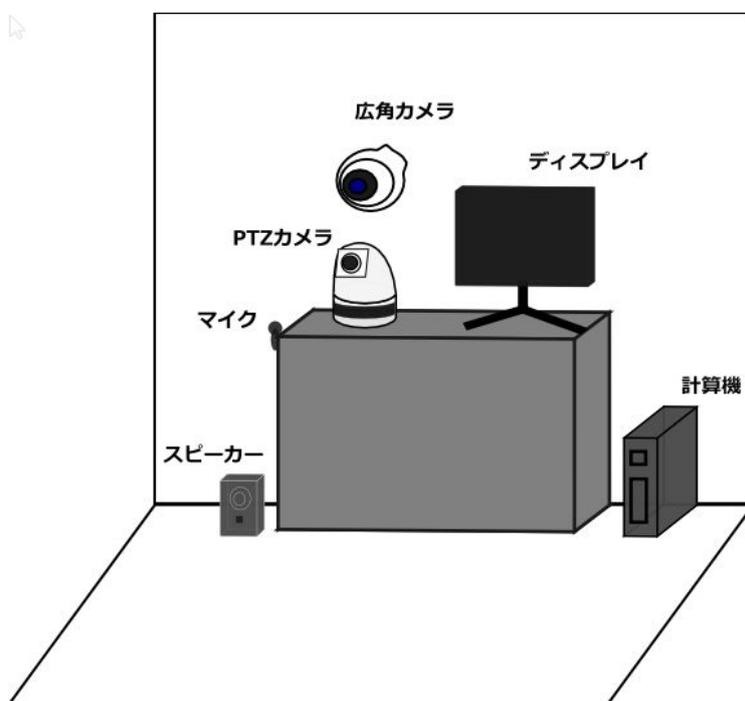


図 4.1: ハードウェア構成のイメージ



図 4.2: 構築したシステムの一部

#### 4.1.1 開発環境

開発言語は C++ , 開発環境は , OS は Windows Vista Home Basic , IDE は VisualStudio2008 を用いて開発を行った . カメラからの画像取得や画像処理を行うためのライブラリとして Intel 社の OpenCV[2] を使用した . システムトレイのアイコンダイアログの表示には MFC(Microsoft Foundation Class) を使用した .

#### 4.1.2 ネットワークカメラ

本研究では AXIS 212 PTZ ネットワークカメラ<sup>1</sup> , AXIS 214 PTZ ネットワークカメラ<sup>2</sup>を使用する . 各ネットワークカメラを図 4.3 に示す . この二つのネットワークカメラはどちらも HTTP リクエストにより , カメラの映像を取得したり , P,T,Z (パン , チルト , ズーム) したりすることが可能である . AXIS 212 PTZ ネットワークカメラと AXIS 214 PTZ ネットワークカメラの技術仕様を表 4.1 に示す .

AXIS 212 PTZ ネットワークカメラは画角が最大 140 °であり , パンチルトすることなく室内のほぼ全体を映すことができる . しかしズーム処理がデジタルで最大 3 倍であるので遠くにある実物体を詳細に映すことは難しい . そのため実物体を映すためにより高解像度のズームが可能である AXIS 214 PTZ ネットワークカメラを使用する .

<sup>1</sup>AXIS: AXIS 212 PTZ <http://www.axiscom.co.jp/prod/212/>

<sup>2</sup>AXIS: AXIS 214 PTZ <http://www.axiscom.co.jp/prod/214/>



AXIS 214 PTZ  
ネットワークカメラ



AXIS 212 PTZ  
ネットワークカメラ

図 4.3: ネットワークカメラ

表 4.1: ネットワークカメラの技術仕様

カメラの型番	AXIS 214 PTZ	AXIS 212 PTZ
水平画角	2.7 ° ~ 48 °	44 ° ~ 140 °
最大解像度 (ピクセル)	704 × 480	640 × 480
最大フレーム数	30fps	30fps
パン	± 170 °	± 70(デジタルパン)
チルト	-30 ° ~ +90 °	± 52(デジタルチルト)
ズーム	18 倍光学 12 倍デジタル	3 倍デジタル

AXIS 214 PTZ ネットワークカメラは、パン、チルトを行うことで広角カメラが映す範囲を撮影することができる。ズーム率が最大 18 倍であるので離れた場所にある実物体でもズームして撮影することが可能である。またカメラ部が可動するためユーザへの視覚的フィードバックを与える効果もある。

以下の URL にアクセスすることでカメラに動作命令を出したり、カメラの映像を所得したりすることができる。

**http://(カメラの IP アドレス)/mjpg/video.mjpg**

カメラからデータストリームを取得する

**http://(カメラの IP アドレス)/axis-cgi/com/ptz.cgi?(パラメータ)=(値)**

パラメータとその値によってカメラの制御を行う。

- (パラメータ) = “center” で (値) = “x,y” に設定した場合、画像上の (x,y) 座標をカメラの視線の参照座標となるようにカメラをパン・チルトさせる。
- (パラメータ) = “pan” で (値) = “p” に設定した場合、p の値に対応させてカメラをパンさせる
- (パラメータ) = “tilt” で (値) = “t” に設定した場合、t の値に対応させてカメラをチルトさせる。

- (パラメータ) = “ zoom ” で (値) = “ z ” に設定した場合，z の値に対応させてカメラをズームさせる．

なお本稿では AXIS 212 PTZ ネットワークカメラを広角カメラ，AXIS 214 PTZ ネットワークカメラを PTZ カメラと記述する．

## 4.2 処理の流れ

システム全体の処理の流れを図 4.4 に示す．システムが開始されるとまず，広角カメラによる手振り動作の認識を行う．広角カメラ映像で手振りが認識されると PTZ カメラの向きを変化させる．次に PTZ カメラにより手振り動作の認識を行う．PTZ カメラで検出した手振り動作の長さによって実物体の登録とデジタルコンテンツの関連付けの処理，または実物体の照合とデジタルコンテンツの参照の処理に分かれる．

処理は手振り動作の認識処理，カメラの操作処理，実物体の認識処理，実物体とデジタルコンテンツの関連付け処理に分かれる．

## 4.3 手振り動作の認識

手振りなどの周期的な動作の認識としては，カラー画像から肌色抽出を行うことで手領域を抽出し，その重心座標の振動成分を評価することで検出する手法などがあげられる [3][4]．しかしこれらの研究は肌色抽出が不可欠であり，照明や肌色の個人差の影響を受けやすいといった問題点がある．また，本研究では実物体を把持していることを想定しているため，肌色抽出は困難である考えられる．そこで本研究では手振り動作の認識に入江ら [5] の手法を実装した．

入江らの手法は，低解像度化した濃淡画像の各画素に対して時間軸方向の FFT(Fast Fourier Transform) を行い，そのスペクトルから手振りの動作の検出を行うものである．濃淡画像を用いることによって，照明条件や肌色の個人差に対してロバストであり，手領域を抽出するといった画像処理は不必要である．また濃淡画像を用いることによって実物体を持った状態でも手振りの動作を認識することが可能である．

本研究ではユーザの位置をカメラから 2m ~ 6m の範囲と想定している．入江らの研究によるとカメラの画角を  $\theta$ [rad]，取得画像の水平画素数を  $a$ [pixel]，最大計測距離を  $L$ [m]，手振りの幅の距離を  $L_H$ [m]，手振りの動作平面と画像平面のなす角度を  $\alpha$ [rad] とすると，取得画像における手の重心移動量  $H$ [pixel] は (4.1) 式であらわされる．手領域と背景の間で濃淡が変化するには手振りの幅  $H$  が低解像度化した画像において 1 画素以上の幅を持たなければならない．

$$H = \frac{aL_H \cos(\alpha)}{2 \tan(\frac{\theta}{2})} \times \frac{1}{L} \quad (4.1)$$

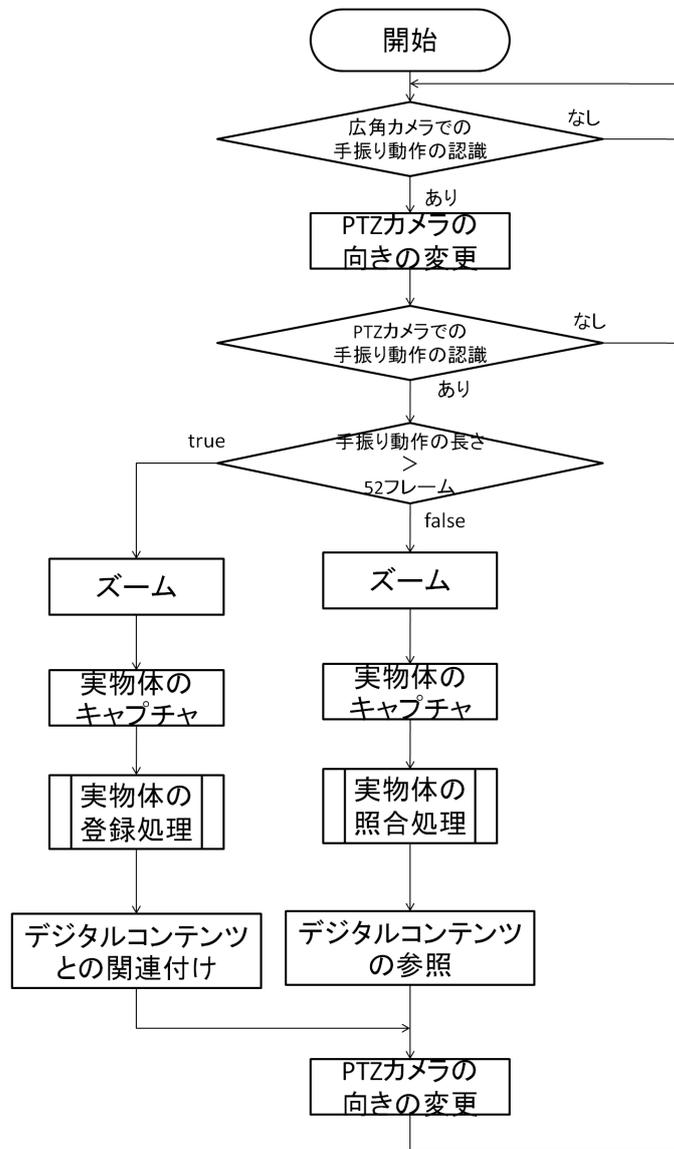


図 4.4: 処理の流れ

### 4.3.1 広角カメラによる周期的な動作の認識

(4.1) 式により解像度を決定した．各パラメータの値は  $\theta = \text{約 } 3 / 4 \text{ [rad]}$  ,  $a = 640 \text{ [pixel]}$  ,  $L = 6 \text{ [m]}$  ,  $L_H = 0.3 \text{ [m]}$  であり ,  $\alpha = 0 \text{ [rad]}$  , として計算すると  $H = 6 \text{ [pixel]}$  となる．そのため広角カメラでは ,  $5 \times 5$  ピクセルを 1 画素とし , 解像度を  $128 \times 96$  とした．図 4.5 に広角カメラからのキャプチャ画像と低解像度化した濃淡画像を示す．



(a) カメラ映像



(b) 低解像度化した濃淡画像

図 4.5: 広角カメラ映像と低解像度化した濃淡画像

本研究ではFFTを行うサンプリング数は16とし、カメラからのフレームレートは20fpsとする．最初に手振り動作を検出した後の10フレームの間、手振り動作を検出し続ける．その間手振りを行っているとして検出された画素すべてを含む最小の矩形を手振り動作をしている範囲とする．検出した手振り動作の範囲の例を図4.6に示す．図4.6の緑色の矩形が検出した手振り動作の範囲である．



図 4.6: 広角カメラでの手振り動作の検出例

### 4.3.2 PTZ カメラによる周期的な動作の認識

広角カメラと同様に(4.1)式により解像度を決定した。各パラメータの値は $\theta = \text{約 } \pi/4[\text{rad}]$ ,  $a = 704[\text{pixel}]$ ,  $L = 6[\text{m}]$ ,  $L_H = 0.3[\text{m}]$ であり, $\alpha = 0[\text{rad}]$ ,として計算すると $H = 39[\text{pixel}]$ となる。より詳細な範囲を検出するため,PTZ カメラでは,20 × 20 ピクセルを1画素とし,20 ピクセル確保することのできない右側部分は4 × 20 ピクセルで1画素とした。解像度を36 × 24 とし行った。図 4.7 にカメラからのキャプチャ画像と低解像度化した濃淡画像を示す。



(a) カメラ映像



(b) 低解像度化した濃淡画像

図 4.7: PTZ カメラ映像と低解像度化した濃淡画像

FFT を行うサンプリング数は16とし,カメラからのフレームレートは20fps とする。最初に手振り動作を検出した後の10フレームの間,手振り動作を検出し続ける。その間手振りを行っているとは検出された画素すべてを含む最小の矩形を手振り動作をしている範囲とする。検出した手振り動作の範囲の例を図 4.8 に示す。図 4.8 の緑色の矩形が検出した手振り動作の範囲である。



図 4.8: PTZ カメラでの手振り動作の検出例

## 4.4 カメラの操作処理

カメラの操作処理は，広角カメラで手振りを検出した位置に PTZ カメラを合わせる広角カメラと PTZ カメラの対応処理，PTZ カメラにより手振りの検出範囲をズームするズーム処理の二つである．

カメラの初期位置は，室内全体を見渡せるように広角カメラを設置し，広角カメラのカメラ画像の中心と PTZ カメラのカメラ画像の中心が同じになるように PTZ カメラを設置する．PTZ カメラのパン，チルト，ズームによる各値の変動は全て内部に記録する．処理が終わったら初期の位置，ズーム率に戻るよう設定する．

### 4.4.1 広角カメラと PTZ カメラの対応処理

広角カメラで手振りの位置を検出したのち PTZ カメラの向きを変化させる．広角カメラでの手振りの認識の後，PTZ カメラによりもう一度手振りの認識を行うので，パン方向の移動は考慮しない．そのため PTZ カメラのパン方向のみ変化させる．

本研究で使用する広角カメラ (AXIS 212 PTZ) の画角は  $140^\circ$  であり，解像度は横 640 ピクセルであるため，広角カメラで検出した周期的な動作の  $x$  方向の中心のピクセルを  $X$  とすると PTZ カメラをパン方向に移動させる角度は (4.2) 式を解くことで得られる．(4.2) 式で得られた移動させる角度がマイナスの場合は左方向に，プラスの場合は右方向にカメラの向きを変える．

$$\text{移動させる角度} = \frac{X}{640} \times 140 - 70 \quad (4.2)$$

### 4.4.2 ズーム処理

PTZ カメラ映像で検出した手振り動作の範囲をもとに，パン，チルト，ズームする値を計算する．検出した手振り動作の範囲の中心のピクセルを求め，そのピクセルに PTZ カメラの中心が移動するようにパン，チルトする値を決める．また AXIS 214 PTZ の解像度は  $704 \times 480$ ，水平画角は  $2.7^\circ \sim 48^\circ$  となっており，カメラのズームレベルは  $1 \sim 9999$  であるため，ズームする値は (4.3) 式を解くことで求めることができる．

$$\text{求めるズーム率} = 48 - \frac{\text{検出範囲の画角}}{48 - 2.7} \times 9999 \quad (4.3)$$

## 4.5 実物体の認識処理

物体の認識には，局所記述子 (local descriptor) によって得られる特徴ベクトルを用いて画像を表現し，特徴ベクトル同士を比較し，検索・識別する手法がある．本研究では局所記述子による特徴ベクトル抽出法の一つである SURF (Speeded Up Robust Features) [6] を用いた．SURF

は画像の回転，スケール変化，照明変化に対してロバストであり，物体の一部に隠れが生じても，見えている部分の特徴量が使用できるといった利点がある。SURF の特徴ベクトルは OpenCV の関数 `cvExtractSURF` を用いることにより抽出している。画像からの SURF の抽出例を図 4.9 に示す。図 4.9 の青い円は SURF 特徴の位置とスケールを表している。

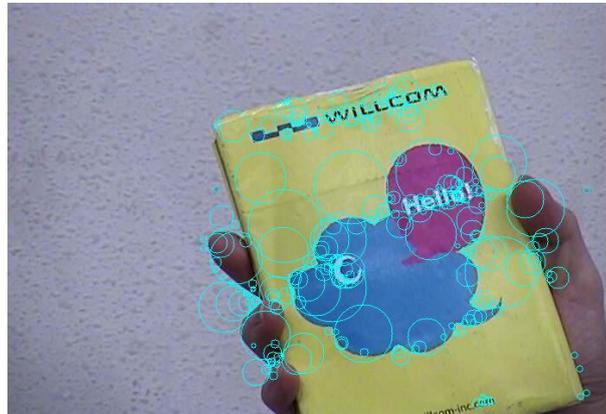


図 4.9: SURF による特徴抽出の例

#### 4.5.1 実物体の登録

実物体の登録時には，実物体をキャプチャした画像に対して SURF を行い，得られた特徴ベクトルを物体 ID とともにデータベースに登録する。また物体 ID と実物体の画像ファイル名のペアも共にデータベースに登録する。データベースに登録されたデータの例を図 4.10 に示す。

物体ID	特徴ベクトル
1	0.01,0.29,0.12,0.30,...
1	0.43,0.05,0.38,0.12,...
1	0.15,0.34,0.18,0.52,...
⋮	⋮
2	0.29,0.21,0.09,0.32,...
2	0.32,0.35,0.24,0.01,...
⋮	⋮

物体ID	物体のファイル名
1	pict1.jpg
2	pict2.jpg
3	pict3.jpg
⋮	⋮

物体IDと特徴ベクトルのペア

物体IDと物体のファイル名のペア

図 4.10: データベースに保存された実物体のデータ例

## 4.5.2 実物体の照合

照合の処理の流れを図 4.11 に示す．照合時にはまずクエリ画像から SURF の特徴ベクトルを抽出する．クエリ画像から抽出した特徴ベクトルをデータベースに登録してある各画像の特徴ベクトルと比較する．クエリ画像と一致している特徴ベクトルが最も多い画像を結果として返す．本研究では，登録したい実物体以外に手や背景も画像として映るため，一致する特徴ベクトルが最も多く，かつ一致する特徴ベクトルの数が入力画像の特徴ベクトルの  $1/3$  以上である画像を一致する画像としている．それ以外の場合は一致する画像がないと判断する．

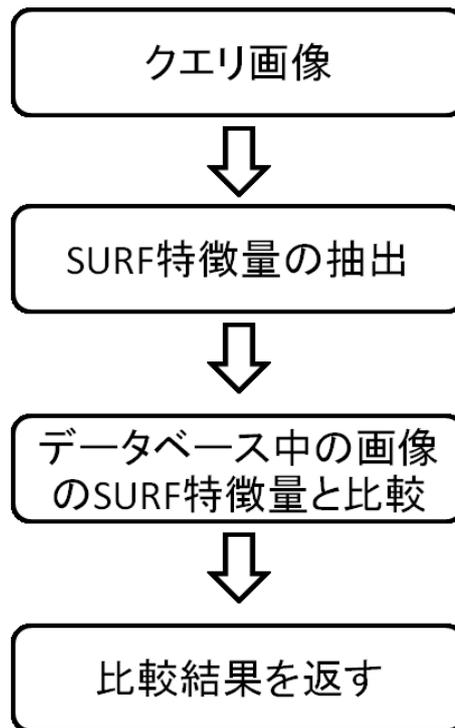


図 4.11: 実物体の照合の処理の流れ

## 4.6 実物体とデジタルコンテンツの関連付け処理

### 4.6.1 実物体と関連付けられたデジタルコンテンツの保存

実物体とデジタルコンテンツの関連付けはデータベースによって管理する．保存するデータは，実物体では実物体の登録の時に用いた物体 ID を使用し，デジタルコンテンツでは，Web ページならば URL, 動画画像や音楽ファイルなどはそのファイルへのパスを使用する．図 4.12 に保存されたデータの例を示す．

物体ID	コンテンツ
1	C:\%video%\video1.avi
2	http://www.iplab.cs.tsukuba.ac.jp
3	C:\%sound%\music1.wav
4	C:\%video%\video3.avi
⋮	⋮

物体IDとデジタルコンテンツのペア

図 4.12: データベースに保存されたコンテンツのデータ例

## 4.6.2 デジタルコンテンツの取得

### URL の取得

Win32API を利用し Internet Explorer で現在開いている Web ページの URL を取得している。

### 既存のファイルのパスの取得

MFC の関数を使用することによりコモンダイアログを表示し、ファイルを選択することにより取得している。

### 動画像・音声の取得

動画像は PTZ カメラの映像を利用する。保存する映像は PTZ カメラのものを保存し、OpenCV の関数を用いて、avi 形式として保存する。音声は、部屋に設置したマイクから取得する。

PTZ カメラ映像で周期的な動作を認識するか一定時間経過すると動画像・音声の取得を終了する。

## 4.6.3 実物体と関連付けられたデジタルコンテンツの参照

コンテンツの参照時には物体認識の結果から得られた物体 ID をキーとして、データベースを検索する。検索結果として返されたデジタルコンテンツに対応した処理を行う。

## 第5章 議論

### 5.1 試用とそこから得られた知見・課題

本研究で開発したシステムの試用を行った。データベースに図 5.1 のように実物体とデジタルコンテンツを関連付けて登録しておいた状態で試用を行った。図 5.2 は物体 ID=3 の実物体を用いて <http://www.iplab.cs.tsukuba.ac.jp/index-j.html> の Web ページを参照している様子である。

実物体画像	物体ID	デジタルコンテンツのパス・URL
	1	C:\%contents%\music1.wav
	2	C:\%contents%\movie.avi
	3	<a href="http://www.iplab.cs.tsukuba.ac.jp/index-j.html">http://www.iplab.cs.tsukuba.ac.jp/index-j.html</a>
	4	C:\%contents%\pict1.jpg

図 5.1: 試用時に用いたデータ

試用を通じて、実物体を用いて関連付けられたデジタルコンテンツを参照することができた。また広角カメラ、PTZ カメラを利用することで、室内のどこにいても実物体を認識させることが可能であった。本システムの利用によりユーザの身の周りの実物体を利用して目的のコンテンツを参照することが可能であった。

またカメラ、マイクにより動画像・音声を録音し実物体に関連付けて保存することもできた。これにより実物体がとっさのメモを保存するモノの役割を果たすことも可能である。

一方実物体をしっかりと認識できないという問題点も発見した。画像をキャプチャする際に実物体以外の部分が多くなってしまい、うまく認識できない場合があった。その原因として、周期運動を行っているときに手以外の腕や肘なども同時に動いてしまうためその範囲が手振りをしてしていると認識されてしまいズーム率が適切でないということが挙げられる。今回の実装ではユーザに与えるフィードバックは音のみであったため、なんらかの方法でズーム範囲



図 5.2: 試用している様子

のフィードバックをユーザに与えることが必要であると考えられる。

## 5.2 発展

### 5.2.1 複数地点での使用

本システムでは一地点での使用を想定しているが、複数地点で使用することができる一つの実物体で、参照したいデジタルコンテンツを様々な場所で使用できる。頻繁に利用するデジタルコンテンツであったり、共有したいコンテンツがある場合に有効であると考えられる。またある地点で動画・音声を記録し、他の場所で参照することが可能であると、同じ実物体を利用したコミュニケーションをとることも可能となると考えられる。

### 5.2.2 複数の関連付けやデジタルコンテンツの操作

本システムでは一つの実物体に一つのデジタルコンテンツしか関連付け、参照することしかできない。実物体に複数のデジタルコンテンツを関連付け、それらを新たなジェスチャなどを通して操作することができるとより柔軟なインタラクションを行えると考えられる。いくつかの Web ページを一つの実物体に関連付けて置くことで一緒に見たい Web ページをまとめて参照することが可能となり、ひとつの物体により多く意味を与えることができると考えられる。

## 第6章 関連研究

### 6.1 身の回りにある実物体を利用した研究

Kawamura らの Ubiquitous Memories[7] は、「記憶を対象に貼りつける」といったコンセプトを提案し、ウェアラブルコンピュータと RFID によるシステムで実現した。対象物に RFID を取り付け、それに「触れる」という行為でユーザの体験映像を保存したり、対象物から体験映像を読みだしたりすることが可能である。本研究とは、実物体に RFID をついたり、ユーザは特別な装置を身につけていない点で異なる。

綾塚らの Ubiquitouslinks[8] は現実世界のオブジェクトから Web の世界の関連する情報のリンクを構築するシステムである。実世界の物体にタグを取り付け、認識装置を取り付けた計算機を近づけることによってそのオブジェクトに関係する Web 上の情報を見ることができる。この研究は、実物体に Web 上の情報を関連付け閲覧するという点で関連するが、実物体にタグを取り付けたり小型の計算機を用いなければいけない点で本研究とは異なる。

Kai らの iCon[9] はユーザの財布やペットボトルなどの日常的な実物体をテーブルトップのコントローラとして利用している。簡単に貼って剥がすマーカーを実物体に貼りカメラで認識することで実物体を補助的な入力装置として用いている。この研究は実物体をテーブルトップのコントローラとして用いている点で本研究とは異なる。

### 6.2 実物体に情報を保存する研究

Brygg らの mediaBlocks[10] は実物体をオンラインデータの物理的なアイコンとみなし、RFID を取り付けた木製のブロックをオンラインデータとリンクさせている。特殊なハードウェアによりブロックの位置を検出ことができ、木製のブロックを操作することにより、ビデオ編集やデータのコピー＆ペーストなどを直観的に操作することが可能である。

美崎らの研究[11]では、個人の見たもの、書いたものなどの体験をデジタル化し、住居を記憶媒体として蓄積させている。住居内のディスプレイに常時スライド表示させ、居住者が受動的に眺めることで個人の記憶を想起させたり、探索活動を支援している。

椎尾らの IconSticker[12] は、アイコンの図柄、名前、コンピュータ可読な情報(バーコード)が印刷された紙にコンピュータ上のアイコンを対応付けることで、コンピュータ画面内だけでなく実世界にアイコンを取り出すことができる。バーコードをバーコードリーダーで読み込むことで、対応するコンピュータのアイコンにアクセスし情報を利用することができる。

Physical Hypermedia[13] は実世界に存在するものをデジタル世界で扱えるようにしたり、デ

デジタルの世界で実現される機能を実世界に提供できるようにする仕組みを提案している。

以上の研究はどれも実世界上の物体にデジタル情報を保存する点で本研究と関連するが、使用する実物体がそれ専用の特別なモノである。本研究ではデジタル情報を保存する物体には身の回りにある実物体を用いる点で異なる。

### 6.3 周期的な動作の認識を活用した研究

入江らや飯田らの研究は、指差しなどのジェスチャにより家電機器を操作する研究である [14][15]。ユーザを特定するため、ユーザの明示的な意思を示すための開始ジェスチャとして手振りなどの周期的な動作を用いている点で本研究と関連する。しかし本研究では、実物体によってデジタルコンテンツを参照することを目的としている点で異なる。

竹内らの研究 [16] では、手振りの速度変化により異なる指示内容を示す「こっち来い」「あっち行け」ジェスチャを認識し、「ページめくり」「ページはらい」動作へと適応している。周期的な動作を直接 PC 操作のインタフェースに利用している点で本研究とは異なる。

浜田らの研究 [17] では、「切る・混ぜる」など、料理中の周期的な動作を検出し、繰り返し動作を検出することで、調理映像における重要映像の抽出を行っている。この研究では、周期的な動作を映像の抽出に利用しており、本研究ではインタラクションの一部として利用している点で異なる。

## 第7章 結論

本研究では、実物体に関連付けられたデジタルコンテンツ参照のためのインタラクション手法を提案し、その手法を実現するシステムを試作した。カメラ映像を通して、ユーザの身の回りにある実物体にデジタルコンテンツを関連付け、室内のどこからでも実物体を利用してデジタルコンテンツを参照することが可能となった。

今後は認識率の向上やシステムの改善を行い、本システムの利用のしやすさなどの客観的な評価を行いたい。

## 謝辞

本研究を進めるにあたり，指導教員である田中二郎教授をはじめ，高橋伸准教授，三末和男准教授，志築文太郎講師には，適切にご指導をいただきました．特に，高橋伸准教授には，ユビキタスチームの担当として研究の初歩から論文執筆に至るまで数多くの助言をいただきました．深く感謝いたします．また，IPLABの皆様，とりわけユビキタスチームの皆様には，ゼミ以外にも数多くのご意見やご指摘をいただきました．厚く御礼申し上げます．最後に，ここまで私を支えてくださいました，家族・友人に心から感謝いたします．

## 参考文献

- [1] Mark Weiser. The computer for the twenty-first century. In Scientific American, pp.94-104, 1991.
- [2] OpenCV. <http://opencvlibrary.sourceforge.net/>
- [3] J. Sherrah and S. Gong. VIGOUR: A System for Tracking and Recognition of Multiple People and their Activities. Proc. of the International Conference on Pattern Recognition, 2000.
- [4] R. Cutler and L.S. Davis. Robust Real-Time Periodic Motion Detection, Analysis, and Applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.8, pp. 781-796, 2000.
- [5] 入江耕太, 梅田和昇. 濃淡値の時系列変化を利用した画像からの手振りの検出. 日本ロボット学会誌, Vol.21, No.8, pp. 923-931, 2003.
- [6] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool. SURF: Speeded Up Robust Features. Computer Vision and Image Understanding(CVIU), Vol. 110, No.3, pp. 346-359, 2008.
- [7] Tatsuyuki Kawamura, Tomohiro Fukuhara, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kido. Ubiquitous Memories: a memory externalization system using physical objects. Personal and Ubiquitous Computing - Memory and Sharing of Experiences, Volume 11 Issue 4, pp. 287-298, 2007.
- [8] 綾塚 祐二, 暦本 純一, 松岡 聡. UbiquitousLinks : 実世界環境に埋め込まれたハイパーメディアリンク. ヒューマンインターフェース研究会報告 67-4, pp. 23-30, 1996.
- [9] Cheng, Kai-Yin and Liang, Rong-Hao and Chen, Bing-Yu and Laing, Rung-Huei and Kuo, Sy-Yen iCon: utilizing everyday objects as additional, auxiliary and instant tabletop controllers. Proceedings of the 28th international conference on Human factors in computing systems(CHI '10). pp. 1155-1164, 2010.
- [10] Brygg Ullmer and Hiroshi Ishii. mediaBlocks: tangible interfaces for online media. In CHI '99: CHI '99 extended abstracts on Human factors in computing systems, pp. 31-32, 1999.
- [11] 美崎薫, 河野恭之. 「記憶する住宅」～55万枚のデジタルスキャン画像の常時スライドショー・ブラウジングによる過去記憶の甦りの実際～, インタラクシオン 2004 論文集, pp. 129-136, 2004.

- [12] 椎尾一郎, 美馬義馬. IconSticker : 紙アイコンによる情報整理. コンピュータソフトウェア Vol.16, NO.6, pp. 24-32, 1999.
- [13] K. Grunbal, J.F. Kristensen, P. Orbak, M.A. Eriksen. "Physical Hypermedia ": Organising Collections of Mixed Physical and Digital Material. Proc. of The 14th conf. onHypertext and Hypermedia(ACM HyperText '03), pp.10-19, 2003.
- [14] 入江耕太, 若村直弘, 梅田和昇. ジェスチャ認識に基づくインテリジェントルームの構築. 日本機械学会論文集 C 編, Vol.73, No.725, pp. 258-265, 2007.
- [15] 飯田直也, 戸澤慶昭, 中村明生. ジェスチャ認識を用いた家電機器操作インタフェースの検討 : 操作開始合図としての手振り動作認識及び操作機器特定のための指差し方向推定. 映像情報メディア学会技術報告 33, pp. 81-84, 2009.
- [16] 竹内麻梨子, 加藤邦人, 山本和彦. 金赤外肌検出法を用いた手振りインタフェースシステムの構築. 映像情報メディア学会技術報告 34, pp. 35-38, 2010.
- [17] 浜田玲子, 佐藤真一, 坂井修一, 田中英彦. 料理映像における繰り返し動作のスポッティング手法. 電子情報通信学会技術報告, パターン認識・メディア理解研究会 PRMU2001, pp. 77-82, 2001.