

話者動作による確実なポインティングのできる プレゼンテーションシステム

若松 翔^{1,a)} 梅村 恭司¹ 岡部 正幸² 三輪 多恵子³

概要: プレゼンテーションにおいてスライドのある部分に聴者の注目を集める方法として、ポインティングデバイスの利用が挙げられる。しかし、ポインティングデバイスは話者が注目箇所を示し続けなければならない、聴者がスライドから目を離していた場合に注目箇所を見逃すと話の流れが分からなくなり、その後の説明についていけなくなる危険性がある。また、話者はスライド上の重要な箇所は説明の際に何度か参照する必要があり、その度にポイントを行うことはプレゼンテーションの負担になってしまう。本研究では、話者のシルエットをスライドの背景として表示し、話者自身がポインティングデバイスとなり発表を行うことができるプレゼンテーションツールにおいて、動作を行うことにより注目箇所をマーキングする方法を示す。

A Presentation System with Reliable Pointing by Presenter's Action

SHO WAKAMATSU^{1,a)} KYOJI UMEMURA¹ MASAYUKI OKABE² TAEKO MIWA³

Abstract: The purpose of pointing device for presentation slides is to attract the attention of listeners to certain place of slides. One problem of pointing device is its unreliability. Once the presenter have moved or removed the pointing mark from the place before listeners actually pay attention the right place, listeners may have difficulty to follow the presentation. On the other hand, it is usually painful for presenters to keep the pointing mark on a topic, while they are referring to the topic. In this paper, we show a new method for showing a mark on presenting slide, using presenter's silhouette and his action. We have implemented reliable marking sign that responds the action of presenter.

1. はじめに

近年、学会や勉強会のネットワーク配信が盛んになっている。このような配信において、スライドが投影されているスクリーンをカメラで撮影したものを配信する場合、ポインティングデバイスを用いると実際に会場にいる人には効果があるが、配信で見ていない人には動画のビットレートによってはスライド及びポインティングの明瞭性が失われ

てしまうという問題がある。話者のコンピュータ出力からのスライドの映像と会場の話者をカメラで撮影した映像を組み合わせて一つの画面で配信するという手法が用いられる場合もあるが、この場合話者の映像を表示させる分スライドの画面が小さくなってしまい、こちらもプレゼンテーションの明瞭性が低下してしまう。

また、このようなネットワーク配信を行わない場合でも、話者のジェスチャーとスクリーンという2つの対象に対して視線移動があり、場合によっては注意が分散しプレゼンテーションの情報を読みそこねる危険性がある。

これらの問題に対し話者の動きのある場所を特定し、それに対応したシルエットを作成した後そのシルエットをスライドの背景に表示するというシステムが提案、開発されている [6]。このシステムでは、人物の動きによる残像がスライドの妨げになってしまい視覚的に見づらいということ、話者がスライドのどこを指しているのかが分かりづら

¹ 豊橋技術科学大学 情報・知能工学系
Department of Computer Science and Engineering, Toyohashi University of Technology

² 豊橋技術科学大学 情報メディア基盤センター
Information & Media Center, Toyohashi University of Technology

³ 豊橋創造大学 経営学部経営学科
Department of Business Administrations, Faculty of Business Administrations, Toyohashi Sozo University

a) wakamatsu@ss.cs.tut.ac.jp

いといった問題があったため、話者のシルエットを常に表示し視覚的变化を少なくすること及び話者のシルエットの一部に色付けを行いスライドのどの部分を指し示しているのかを明確にするという処理を施し、シルエットの安定化とポインティング機能の追加を行った。

しかしこのシステムでは主張部分のポイントは話者が指し示している間しか行われないため、聴者がスライドから目を離している際にポイントがあった場合、聴者は重要箇所を見過ごしその後の説明についていけなくなるという危険性がある。また、話者のほうも重要箇所は何度か参照する必要があり、その度にスライド上でポイントをしなければならずプレゼンテーションの負担となってしまう。これはこのシステムに限らずレーザーポインターを使用する場合でも同様であり、多くのプレゼンテーションにみられる問題である。

そこで話者がスライド中の重要箇所に対し手をセンサ (Kinect) に向けて押し出す動作 (プッシュ動作) を行うことにより、その部分にスタンプが出現しスライド上にマーキングを行うことができる機能の実装を行った。これにより聴者はスライド中の重要箇所を見過ごすことがなくなり、話者に対してはプレゼンテーション時の負担を下げる事が可能となる。

2. システムの概要

本システムは Microsoft 社の Kinect を用いたプレゼンテーションツールとして実装を行なっている。プレゼンテーションツールとしては、キーボードまたはマウスの操作によってスライドを順番または逆順に表示していく最小限の実装であり、そのスライド上に Kinect で取得したユーザの深度情報のシルエットを常に表示するというものである (図 1)。

目的

- 既存の問題点を解決し、話者のシルエットをスライドの背景として表示するプレゼンテーションシステムを開発する



図 1 話者シルエットによるスライド表示

フレームレートは毎秒 30 フレーム、フレームの大きさは幅 640 ドット、高さ 480 ドットであり、Kinect が動作する環境であれば容易に使用することができる。また、話者が現在スライド上のどの部分を説明しているかを明確にする

ために話者のシルエットの一部を色付けするポインティング機能がある (図 2)。シルエットの色付けは話者とセンサの距離に応じて行われ、単純にセンサに近づくとその部分のシルエットが色付けされる。

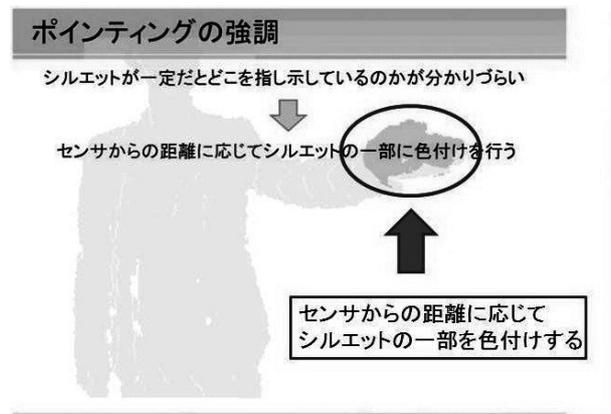


図 2 話者シルエットによるポインティング

3. シルエットの安定化

システムの実装において、Kinect での話者の認識及び背景マスクをおこなったが、この際に関係のないところまでシルエットが描画されてしまい話者のシルエットと背景がちらつくという問題が発生した。このままではこのちらつきにより、スライドとシルエットよりもそちらに注意が分散してしまいプレゼンテーションの明瞭性が下がってしまう恐れがある。

これを解決するために、「描画するシルエットの過去 5 回分の pixel 値を各 pixel ごとに保存しておき、対象の pixel が過去 5 回全てにおいて話者が認識されていれば描画する」という方法を取った。

これは、シルエットや背景のちらつきが影色と白色が交互に描画されることにより発生するためである。具体的には図 3 のようなコードで判定を行なった。図 3 において、 $prevpixel[i]$ がその pixel における話者の認識変数である。 $1prevpixel$ が 1 つ前のユーザ認識フラグ、 $2prevpixel$ が 2 つ前と続いていき、5 つ前 ($5prevpixel$) の話者認識フラグまでを見る。この 5 つで AND を取り、全て話者が認識されていたとしたらシルエットが描画される。それ以外の場合は、シルエットは描画されない。

結果として、話者及び背景のちらつきをほとんど出さないようにシステムを改善することができた。また条件は、話者が認識されている、されていない、のどちらかなので話者が認識されているという条件が 5 回連続で起こることは $1/32$ と有意水準 5% で有意である。条件を「過去 8 回分の pixel 値を保存しておき 8 回連続でその pixel において話者が認識される」とすれば、有意水準 1% で有意となるのだが、8 回連続で対象の pixel において話者が認識されると

```

if (human) {
  if (1prevpixel[i] && 2prevpixel[i]
    && 3prevpixel[i] && 4prevpixel[i]
    && 5prevpixel[i] ) {
    //gray
    outputPixel[i] = shadow_color;
  }
  else{
    //white
    outputPixel[i] = white_color;
  }
}
}

```

図 3 シルエット作成コード

すると、その分シルエットの描画も制限されてしまうため、話者の動きにおける応答が遅くなってしまいます。これにより、手や頭が描画されなくなってしまうため過去5回分とした。過去5回分の場合でも話者の動きにおける応答は遅くなり、速く動かすと手や頭などは多少描画されなくなってしまうが、手を動かしていることや頭が動いているといった話者の動作は十分確認でき、プレゼンテーションの明瞭性に関わる程でもないため良しとした。

4. スタンプ機能

4.1 話者のポインティングにおける要求分析

レーザーポインターのようなポインティングデバイスでスライド上の注目箇所を指し示す場合、聴者の注目を集めることができるがこれは話者が注目箇所を指し示している間のみ有効であり、一瞬でも目を離していると聴者は注目箇所を見過ごす危険性がある。

これは図2におけるシルエットの色付けによるポインティングでも同様のことが言え、話者がポインティングを行なっている際には聴者の注目をポインティング部分に向けることができるが、スライドを見ていなかった聴者はどこがポインティングされたのか分からないため重要な箇所を見逃したことによりこの先の説明についていけなくなる危険性がある。

そのため話者がスライド上の注目箇所を指定することにより、その部分がマーキングされるような機能があれば、聴者はスライドから目を離していた際にも瞬時にマーキング箇所が重要であることが理解でき、話の流れについていけなくなるということもなくなる。また、話者にとってもスライドでの重要な箇所は何度か参照するため、いちいちその箇所をポインティングする必要がなくなるといった利点もある。

4.2 マーキング方法における仕様検討

スライド中での重要箇所を説明する際に、話者は指定の箇所に対し何らかのアクションを起こす必要がある。そ

の際に、聴者が指定箇所が重要であることを話者の動作によっても理解できるように、マーキング方法は話者自身の身体的な動作によることが望ましい。

そこで、話者がスライド上の指定箇所を手でプッシュすることによりマーキングできるような手法を考えた。マーキングはスライドへの影響も考え、単純に円形のスタンプイメージがスライド上に表示されるようにした(図4)。ス

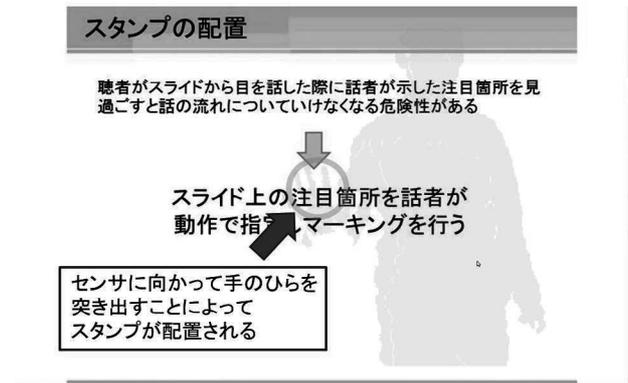


図 4 プッシュ動作によるスタンプの配置

スタンプが複数配置されると聴者の視線が分散してしまい、混乱を招く恐れがあるためスライド上には一つだけしか配置できないようにした。スタンプが画面上に表示されている間にもう一度プッシュ動作を行うと現在表示されているスタンプは削除され、プッシュ動作を行った位置に新たなスタンプが配置される。また、一度配置したスタンプがいつまでも残っているとスライド自体の妨げとなるため、スタンプは徐々に透明になっていき最終的には画面上から消えるといった挙動を行う(図5)。

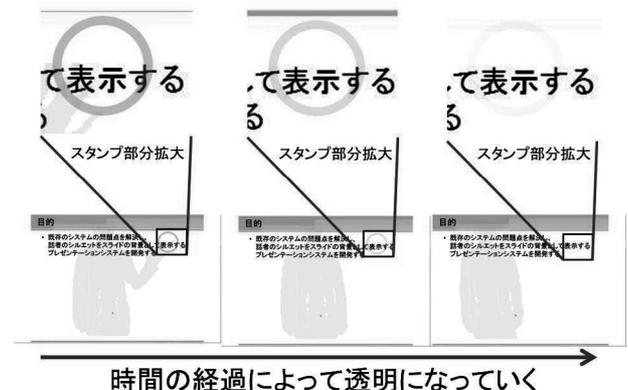


図 5 時間経過によるスタンプの透明度変化

4.3 スタンプ機能の実装

プッシュ動作によるスタンプの配置は、シルエットの深度情報をフレーム毎に全画素1ピクセルずつ保存しておき、あるピクセルの過去2フレームを比較した際に深度の差が閾値を超えているというものが一定以上のピクセルで確認できた場合、画面上にスタンプを配置するという手法で実装を行った。

これにより、話者がスライドの指定箇所をプッシュすることによりそこにスタンプが配置されるようになる。スタンプの表示座標はスタンプを配置すると判断したピクセルのうち、最も深度が小さい(すなわちセンサに近い)座標を中心座標とし、設置される。また、スタンプを設置するとスタンプにタイマーが働き、10秒かけてアルファ値が1から0へ徐々に変化していき、アルファ値が0になるとスタンプは完全に見えなくなり削除される。

5. 評価

実装したスタンプ機能が有用であるかどうかを確かめるために、スタンプ機能の on/off を設定できるようにし、その効果についての有用性の評価を行った。

評価内容は以下の3パターンである。

- A:話者のシルエットと色付けとスタンプの表示
- B:話者のシルエットと色付けの表示
- C:話者のシルエットの表示

以上の条件でそれぞれプレゼンを行い、被験者に対してA(話者のシルエットと色付けとスタンプの表示)とB(話者のシルエットと色付け)とA(話者のシルエットと色付けとスタンプの表示)とC(話者のシルエットの表示)の2つを比較し、伝わりやすい条件はどちらかという設問に対し、どちらが伝わりやすいかを回答してもらった。またスタンプ機能の効果とは関係しないがシステムの参考情報としてB(話者のシルエットと色付け)とC(話者のシルエットの表示)の比較も回答してもらった。

実際のアンケートでは、「条件Xが伝わりやすい」、「迷うが条件Xが伝わりやすい」のように程度を記入してもらい、迷った場合でもどれかの選択肢を選択することを強制した。集計においては、「迷いがある」ということは無視し、選んだ条件がどれかを判定するようにした。結果の分析には片側検定の符号検定を使用する。

評価実験は2度行ったが実験状況が異なるため、それぞれ評価実験1、評価実験2と分けて表記する。評価実験1は実験時間45分、被験者は19人であり、こちらはシルエットの安定化ができていないシステムを使用したため、若干シルエットにちらつきが生じ、スタンプの設置もしづらくなってしまっている。そのため、評価実験2ではシルエットの安定化処理を行い、実験時間30分、被験者は14人である。

また、評価実験1の結果からスライドの背景にシルエッ

トが表示されることは初めて見るものにとっては少なからず違和感が生じることが分かったので、聴者に慣れさせる準備時間として90分間本システムを使用しての講義を行った後、実験を行った。

5.1 評価実験1(実験時間: 45分, 被験者数 :19人, シルエットの安定性が低い)の結果

評価結果を表1に示す。表1より、AとBの比較はB(話者のシルエットと色付けの表示)と比べてA(話者のシルエットと色付けとスタンプ表示)の方が伝わりにくいが6、伝わりやすいが13とA(話者のシルエットと色付けとスタンプ表示)が優れているという結果であり、p値は0.08353で危険率5%では有意と言えない。

これは、シルエットの安定性の低さからスタンプの設置が上手くいかず、操作に手間取ったことと設置の機会が減ってしまったことから色付けによるポインティングとの違いを聴者に感じてもらえなかったことが原因であると推測できる。

次に、AとCの比較はC(話者のシルエットの表示)と比べてA(話者のシルエットと色付けとスタンプ表示)の方が伝わりにくいが3、伝わりやすいが16とA(話者のシルエットと色付けとスタンプ表示)が優れているという結果であり、p値は0.002213で危険率1%と考えた時に有意であり、話者のシルエットのみの表示よりも色付けとスタンプ機能がある方が有効であることが分かった。

また、直接のスタンプ機能との比較ではないため表には記載していないが、BとCの比較はB(話者のシルエットと色付けの表示)と比べてC(話者のシルエットの表示)の方が伝わりにくいが14、伝わりやすいが5とB(話者のシルエットと色付けの表示)が優れているという結果であり、p値は0.03178と危険率5%と考えた時に有意である。

先ほどのA(話者のシルエットと色付けとスタンプ表示)とC(話者のシルエットの表示)との比較の方が良い結果なので、これよりシルエットの安定性が低い場合においてもスタンプのマーカとしての効果があるということが分かった。

5.2 評価実験2(準備時間: 90分 + 実験時間: 30分, 被験者数 :14人)の結果

評価結果を表2に示す。表2より、AとBの比較はB(話者のシルエットと色付けの表示)と比べてA(話者のシルエットと色付けとスタンプ表示)の方が伝わりにくいが1、伝わりやすいが13とA(話者のシルエットと色付けとスタンプ表示)が優れているという結果であり、p値は0.0009155で危険率1%と考えた時に有意であり、シルエットの色付けによるポインティング以上にスタンプのマーカとしての機能が情報伝達において有意に働いていることが伺える。

次に、AとCの比較はC(話者のシルエットの表示)と比

表 1 スタンプ機能の評価結果 1(実験時間: 45 分, 被験者数 :19 人, シルエットの安定性が低い)

	伝わりにくい	伝わりやすい	p 値
B と比べて A は	6	13	0.08353
C と比べて A は	3	16	0.002213

A:話者のシルエットと色付けとスタンプ表示

B:話者のシルエットと色付けの表示

C:話者のシルエットの表示

表 2 スタンプ機能の評価結果 2(準備時間: 90 分 + 実験時間: 30 分, 被験者数 :14 人)

	伝わりにくい	伝わりやすい	p 値
B と比べて A は	1	13	0.0009155
C と比べて A は	1	13	0.0009155

A:話者のシルエットと色付けとスタンプ表示

B:話者のシルエットと色付けの表示

C:話者のシルエットの表示

べて A(話者のシルエットと色付けとスタンプ表示)の方が伝わりにくいが 1, 伝わりやすいが 13 とこちらも A(話者のシルエットと色付けとスタンプ表示)が優れているという結果であり, p 値は 0.0009155 で危険率 1%と考えた時に有意であり, 話者のシルエットのみの表示よりも色付けとスタンプ機能がある方が有効であることが分かった。

また, スタンプ機能の効果には関係しないがシステムの参考情報として B と C の比較は B(話者のシルエットと色付けの表示)と比べて C(話者のシルエットの表示)の方が伝わりにくいが 14, 伝わりやすいが 0 と B(話者のシルエットと色付けの表示)が優れているという結果であり, p 値は 6.104×10^{-5} と危険率 1%と考えた時に有意であり, 検証済みではあるがシルエットの色付け機能が有効であることが分かった。

5.3 評価実験のまとめ

評価実験 1 と評価実験 2 の評価をまとめると, シルエットの安定性にかかわらず実装したスタンプ機能が有用であることが分かった。

また評価実験 1 より評価実験 2 の結果がシルエットが安定していたことを差し引いたとしても良いのは, 評価実験 2 で設けた 90 分の準備時間により, 聴者の本システムへの慣れによる効果があったからであると思われる。

スライドの背景にシルエットが表示されることは初めて見るものにとっては少なからず違和感が生じる。この違和感を払拭するためにはある程度の慣れが必要であり, 準備時間を設けた評価実験 2 ではシルエットによる違和感の払拭ができたと推測できる。

6. 今後の課題

6.1 マーキング動作に対する議論

スタンプを設置する動作は話者がスライド上の指定箇所

を手でプッシュするというものであるが, このプッシュ動作は過去 2 フレーム間での比較で行なっているため, ある程度勢いがないと画面上にスタンプは設置されない。これは話者の通常の動作中にスタンプの誤設置が起こらないようにするための処置であるが, これによりプッシュしてもスタンプが設置されないということも発生してしまった。

そのため, このような設置動作の判断をどのように切り分けるかが今後の課題である。今回は話者の動きの変化などは無視してしまっているため, このような動作を特定しマーキング機能の向上を図りたい。

6.2 スタンプイメージの議論

評価実験にて同時に実施した自由記述においてはスタンプの形や色が変わらたらいいといった意見が多く見られた。そのため, スタンプの形や色を用途によって変更できれば聴者の更なる注目を集められると考えられる。

今回はスライドへの影響も考え単純に丸の形にしたが, 形の面では星型やパンクで破裂した形, 花などの放射状の形状はよく目立ち, 色の面では青, 紫のような寒色よりも赤, 黄色などのような暖色が注目を集めることができる [5] ため, デザインの面における効果も今後取り入れていきたい。

6.3 スタンプ表示位置における議論

スタンプはスタンプを配置すると判断したピクセルのうち, 最も深度が小さい(すなわちセンサに近い)座標を中心座標とし設置されるのだが, これは手を一番突き出したポイントにスタンプが表示されるという話者へのイメージの単純化のためにこのような仕様にした。

しかし, これは評価実験にて同時に実施した自由記述においても多くみられたのだが, スタンプが適切な場所に表示されないことが度々起こってしまうという結果となった。そのため, 話者が意図したポイントに設置するには何度か

やり直さなければならず負担になってしまい、尚且つ聴者においても動作が気になってしまう。このスタンプの表示位置の正確さがこのシステムの評価に直接結びつくものであるため、今後は話者が意図したポイントにスタンプをより正確におけるようにできる手法を考えていきたい。

7. 関連研究

本システムと同様に、作業空間と人物画像の融合を行ったシステムとしては、ClearBoard[2]がある。本研究はプレゼンテーションによる注目箇所を強調することに焦点をおいたため、ClearBoardとは問題設定が異なるが、遠隔の相手に自身の行動を通じて意思を伝えることができるという点において本システムと共通する。

話者の動作と話者が現在説明している箇所が同時に分かるという点においては、ペンの影を用いてスライドを指示する方法 [4] と共通する。web 配信に利用できるという点において、本システムでは話者自身の動きを見ることができ、指定箇所を動作によりマーキングできるという利点があるが、こちらはペンの影による多彩な指示ができるため細かい表現が可能である。

話者の動作をそのまま反映させるのではなく、タッチスクリーンデバイスを用いてスライド内の注目部分を指定することができるシステム [1] では、話者の身振り手振りのような身体的動作と直感的なデバイス操作により、体による伝達を發揮したプレゼンが可能である。デバイスを通してスポットライトで一部分を照らすような演出も行えるため、本システムと同様に聴者がスライドから目を離し、再度スクリーンを見た際に話者が説明している所が分かるということが可能である。

また、同様にスクリーン前でプレゼンテーションを行うことを支援するシステムである MAEDE[3] ではジェスチャ操作によりスクリーン上のオブジェクトの拡大ができるため、聴者が目を離した際にもどの部分が現在説明されているかが分かるようになっている。

プレゼンテーション中にスライドの一部分を強調したり、スライドの注目点にマーキングができるような聴者の注目を一部分に集めることができるツールとしては、Microsoft の ZoomIt[8] やワイヤレスポインティングパッドを利用した UPIC CoCo[7] などがあり上手く活用すれば効果的なプレゼンテーションが可能になるが、話者の負担が大きくなってしまう。

8. まとめ

本研究では、話者のシルエットをスライドの背景として表示し、話者自身がポインティングデバイスとなるプレゼンテーションツールにおいて、動作により注目箇所をマーキングする方法を示した。話者の指定した注目点に対する聴者の見過ごしをなくすため、スライド上をマーキングで

きる手法を提案し、評価実験によりこれが有効であることが分かった。また、シルエットを安定なものへと処理することにより、既存のシステムにおける視覚的な刺激に対する問題を解決した。

参考文献

- [1] 安保翔太, 山本倫也, 渡辺富夫, タッチスクリーンデバイスを用いた身体拡張型プレゼンテーション支援システムの開発, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会, pp.161-162,(2010).
- [2] 小林稔, 石井裕, ClearBoard-2 における協同作業空間と会話空間のシームレスな融合, 情報処理学会 情報処理学会研究報告, pp.43-50,(1993)
- [3] 前田晴己, 黒澤祐也, 栗原一貴, 宮下芳明, MAEDE:スクリーン前でのプレゼンテーションスタイル, 日本ソフトウェア科学会研究会 WISS2011, pp.164-166,(2011).
- [4] 村田雄一, 志築文太郎, 田中二郎, Shadowgraph: ペンの影を用いて OHP 風の指示ができるプレゼンテーションツール, 第 16 回インタラクティブシステムとソフトウェアに関するワークショップ論文集, 日本ソフトウェア科学会, pp.73-78,(2008).
- [5] 武田瑛夢, やさしいデザイン, エムディエヌコーポレーション, pp.28,106,(2007).
- [6] 梅村恭司, 梅村真由, Kuroko: 話者シルエットを活用するプレゼンツール, 情報処理学会 インタラクション 2012 , pp.281-286,(2012).
- [7] UPIC CoCo
<http://www.plus-vision.com/jp/product/upic/coco-index.html>
- [8] ZoomIt
<http://technet.microsoft.com/ja-jp/sysinternals/bb897434.aspx>