

仮想空間で実世界のアウェアネス情報を取り込んだ効率よい コミュニケーションを実現するためのフレームワークの提案

松田 智 柴田 直樹[†] 安本 慶一 伊藤 実

奈良先端科学技術大学院大学情報科学研究科[†] 滋賀大学情報管理学科

遠隔地の多数のユーザが参加するリアルタイムコミュニケーションシステムは、CSCW (Computer Supported Cooperative Work) などの分野で重要である。多人数参加型コミュニケーションツールとしてはテレビ会議システムが広く利用されており、表情やしぐさなどのアウェアネス情報を伝達できる点で優れている。しかしテレビ会議では、空間におけるユーザの位置関係に基づく情報伝達ができないという問題があり、多人数が参加している環境で、サブグループを作ってコミュニケーションするといったことを直感的な操作で実現することは難しい。本稿ではオンラインゲームで用いられている共有仮想空間の手法を活用し、位置関係に基づいたグループ形成が可能なコミュニケーション環境を実現するためのフレームワークを提案する。提案するフレームワークでは、テレビ会議システムのようにユーザの音声・映像データを伝送することで、現実世界のアウェアネス情報を伝達することを可能とする。仮想空間での位置関係に応じてどのようにアウェアネス情報の通信を制御すれば現実世界のような自然なコミュニケーションに近づくかを考察し、音声や映像などメディアの特性を活かした効率的なコミュニケーションを実現する。

Framework for Supporting Awareness in Virtual Space Communication

Tomo Matsuda Naoki Shibata[†] Keiichi Yasumoto Minoru Ito
Nara Institute of Science and Technology [†] Shiga University

Remote multi-user communication systems play important role especially in CSCW (Computer Supported Cooperative Work). As such tools, video conferencing systems are widely used because awareness information such as facial expression and gestures can be transported. Problem of video conferencing systems is lack of mechanism to transport the information based on spatial relationship of users. When many users participate in video conference, it is difficult to form sub-group by natural manipulation and to communicate with specific person. In this paper, we propose a framework which enables remote users to form group based on spatial relationship by using shared virtual space which has been used as network game framework. Proposed framework can transport awareness information of real world by capturing and transferring user's audio and video data similarly to video conferencing systems. We discuss about how transportation of awareness information should be controlled to model natural communication in real world, and realize efficient communication with the advantage of audio and video media.

1 はじめに

多数のユーザがネットワークを介して実時間でコミュニケーションするためのシステムは、遠隔地にいる人々が協調して仕事や研究を進めるのに役立っている。従来インターネット上でコミュニケーションする為の媒体はテキストを主としてきたが、近年のネットワークの大容量・高速化に伴い、音声によるボイスチャット、さらに音声と映像によるテレビ会議も広く利用されるようになってきた。

テキストだけでなく音声や映像が利用されるのは、コミュニケーションにおいてアウェアネスが重要な役割を果たしているからである。アウェアネスとは人間が何ら

かの行為をするときにそれを取り巻く環境や他者について直感的に把握する知識のことであり、本稿ではコミュニケーションの場面でのアウェアネスを扱う。現実世界の対面コミュニケーションでは、言語以外の情報、例えば相手の表情、視線、身振り、手振り、姿勢、声の性状(声色や抑揚)、さらに匂いや体温などの情報が無意識に利用されている。これらの非言語情報を基に相手の存在、意図、関心、感情などを判断し、コミュニケーションにフィードバックすることで、円滑に意思疎通することが可能となっている [1]。

テレビ会議は、音声に加えて映像を伝送することにより、表情や身振りなどのアウェアネス情報を伝達できる

点が優れている。しかしテレビ会議は参加者の映像が画面上に平面的に配置されるため、空間的な位置に関する情報を伝達できないという問題がある。現実世界の対面コミュニケーションでは相手との関係の度合いに応じて距離や空間を使い分けている [2]。また多人数が集まっているとき、その中の特定の人やグループと会話したい場合に、近づいたり離れたりすることで会話する相手を限定している。このような機能は多数の人が協調して作業を進める CSCW (Computer Supported Cooperative Work) などの分野では必須である。

ネットワーク上で位置関係に応じて会話する相手を選ぶことができるシステムとしてオンラインゲームがある。インターネットに接続した多数のプレーヤがネットワーク上で仮想的な世界を共有し、仮想世界の中で出会った他のプレーヤと会話することができる。しかしオンラインゲームの目的は、現実とは異なる別世界を楽しむことであり、現実世界の情報、例えばユーザの映像などは排除されている。コミュニケーションの媒体はテキストのみに限定していることが多い。

地理的に離れた現実世界の空間を仮想的に統合し CSCW 環境を構築する試みとして Tele-immersion (臨場感通信) の研究も行われている。TEEVEE (Tele-immersion Environment for EVERYbody)[3] は、遠隔地のユーザに対して、各ユーザの視点で互いの 3D 実写映像を実時間で再現し、協調作業環境を実現している。しかし Tele-immersion 環境を構築するには 3D カメラや広帯域のネットワーク環境など特殊かつ高価な装置や設備を必要とし、普及性の面で問題がある。

我々は、(1) テレビ会議のようにユーザの音声と映像を伝えることで多くのアウェアネス情報を伝達することができる、(2) ユーザ間の位置関係に基づいたコミュニケーションをすることができる、(3) 特殊かつ高価な装置を使用することなく、一般ユーザにも普及している PC、ヘッドセット、Web カメラ、インターネット接続環境のみで実現できるコミュニケーションツールが必要だと考える。本稿では、上記 (1) - (3) を実現するためのフレームワークを提案する。

2 関連研究

遠隔地の多数のユーザがリアルタイムでコミュニケーションを行うためのシステムで従来から問題になっていたのは、多数のユーザが参加している環境で、参加者間のインタラクションをどのように制御すれば現実世界の自然なコミュニケーションに近づけることができるかという点である。

各ユーザに対して、システムに参加している全てのユーザに関する情報が等量に与えられると、会話するために必要な情報が他の情報によって干渉を受けて、コミュニ

ケーションが困難になる。またコンピュータの処理能力やネットワーク帯域幅の制約により、全てのユーザが音声データや映像データを伝送するのは不可能な場合がある。効率的なコミュニケーションを実現するには、そのために必要な情報のみを伝達し合うようにユーザ間のインタラクションを制御する必要がある。そこで現実世界でのコミュニケーションのように自然な方法でこれを実現するために、空間に基づいたモデルが提案されてきた。

DIVE (Distributed Interactive Virtual Environment) では、「オーラ」、「フォーカス」、「ニンバス」という概念が提唱された [4]。仮想空間において、各オブジェクトを中心とする一定の空間的な範囲をそのオブジェクトの「オーラ」と定義する。オーラはオブジェクトが移動するとそれに伴って移動する。そして自分のオーラと他のオブジェクトのオーラが交差したとき、オブジェクト間でのインタラクションが可能となる。図 1 において A と B はインタラクション可能であるが、C はインタラクションできない。インタラクションするための媒体は、テキスト、音声、映像など複数あるので、媒体に応じて異なるオーラの距離を設定することができる。例えば相手の姿を見ることは出来るが、声は聞こえないといった現実世界に近い自然な状態を作り出すことができる。オーラを人に適用して他者のオーラと重なると対話を可能にする、オーラを物やサービスに適用して、物のオーラと重なるとサービスを利用可能にするような使い方がされている。

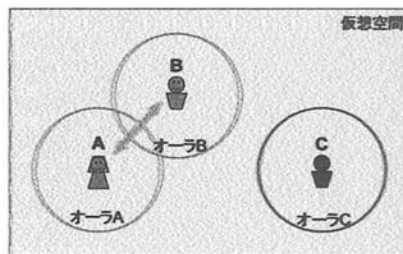


図 1: オーラの概念図

オーラがオブジェクト間の対称的な関係を前提とするのに対し、フォーカスとニンバスは非対称的な関係を前提としている。フォーカスはユーザの可視範囲であり、ニンバスは対象の被可視範囲である。そして自分のフォーカスが対象のニンバスに重なる場合にのみ見ることができる。図 2 で A は B を見ることができるが、B は A を見ることができない。フォーカスとニンバスはオブジェクト間のインタラクションのレベルを制御する役割も果たす。ユーザのフォーカスの、より内部にあるオブジェクトを、より詳しく見ることができる。またオブジェクトのニンバスの、より内部にユーザがいるとき、より詳

しく見ることができる。

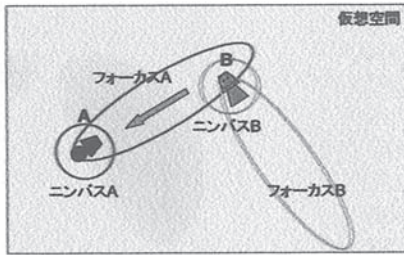


図 2: フォーカスとニンバスの概念図

オーラ、フォーカス、ニンバスの概念は、仮想空間を利用したコミュニケーションシステムで様々に応用され利用されている。例えばカクテルパーティ効果を伴う 3D 音響効果を実現した音声チャットシステムである RAVITAS (Realistic Voice Chat Framework For Cooperative Virtual Space) がある [5]。RAVITAS では音源の音量とそこからの距離に応じて被可聴範囲が決定される。これはニンバスの考え方に近い。また一定の距離内にいるユーザが会話グループを形成し、会話グループのメンバーの音声はより鮮明に聞こえるように制御することでカクテルパーティ効果を実現している。これはオーラの考え方に基づいている。

3 提案するフレームワーク

本章では、フレームワークの目的とそれが実現する機能について述べる。

3.1 フレームワークの目的

現在、多人数参加型のリアルタイムコミュニケーションシステムとしてテレビ会議システムが普及しているが、遠隔地にいるユーザ間の協調作業で広く利用されるようになるとその問題点が明らかになってきた。

協調作業では表 1 のような作業環境や共同作業についてのアウェアネスが必要となる [6][7]。しかしテレビ会議システムではビデオカメラの技術的な制約によりこれらの情報の多くが欠如している。人間の視界と比較するとビデオカメラの視野範囲は狭いので相手の環境について十分な情報を得ることができない。また遠隔地のビデオカメラを思い通りに動かすのは難しく現実世界のように自由に移動しながら情報を得ることもできない [8]。

こうした制約が特に問題になるのは、多人数が参加している環境でその中の特定の人と話をしたい場合や、サブグループに分かれて作業したい場合である。現実世界では「近づく」、「離れる」といった身体性に根ざした直感的な方法でこれを行っているが、テレビ会議システムでは相手を明示的に指定するなどの煩雑な操作が必要になる。

表 1: 協調作業のアウェアネス

Who	誰がいるのか？
	誰がそれをしているのか？
What	何をしているのか？
	何を目的にしているのか？
	何に対して作業しているのか？
Where	どこで作業しているのか？
	どこを見ているのか？
	どこまで見ることができるのか？
	どこに手が届くのか？

一方、オンラインゲームではユーザが仮想空間を共有し、その中を自由に移動しながらそこで出会った他のユーザと会話することができる。しかしテレビ会議のように映像を通じて互いの表情を見ることができない。コミュニケーションにおいて、表情は他人に態度や感情を伝える表現として最も影響力がある重要な情報である [1]。

そこで、(1) オンラインゲームのように仮想空間を共有することにより空間での位置関係に基づく柔軟なグループ編成を可能とし、(2) テレビ会議のように映像と音声を送送することで互いの表情や声の性状などのアウェアネス情報を伝えることができるコミュニケーション環境を実現する。またより多くの人が利用できるように、(3) 一般ユーザに普及している PC、ヘッドセット、Web カメラ、インターネット接続環境でこれを構成する。

フレームワークが対象とするアプリケーションは、コミュニケーションが重視される分野での CSCW、例えば知識創造の分野などを想定する。知識創造ではブレインストーミングなどの手法が活用されており、柔軟にグループを作ったり、表情や身振りを交えた豊かな表現が可能なコミュニケーション環境が望まれている。

また、e-ラーニングでの活用も想定する。近年、教育分野では協調学習が目立っておりグループ活動を中心とする学習形態への取り組みが盛んである [9]。このようなクラスの運営を e-ラーニングでスムーズに行うには、直感的に柔軟なグループ編成ができるコミュニケーション環境が必要である。

3.2 フレームワークの概要

フレームワークは、(1) 仮想空間共有機能、(2) 音声会話機能、(3) ユーザ映像表示機能により構成される。

仮想空間共有機能は、ユーザが三次元の仮想空間を共有し、その中を自由に移動することができる機能である。音声会話機能は、仮想空間の中で出会った他のユーザと音声による会話をすることができる機能である。映像通信機能は、Web カメラでキャプチャしたユーザの実写映像を表示することができる機能である。ユーザに提示さ

れる画面イメージは図3のようになる。



図 3: ユーザ画面のイメージ図

以上の機能は、有機的に連携させることで自然なコミュニケーションに近づけることができる。そこで仮想空間共有機能が提供する空間におけるユーザ間の位置関係に基づいて、音声会話機能とユーザ映像表示機能が提供するユーザ間のインタラクションを制御するモデルを考える。

現実世界のコミュニケーションにおいて他者を認知する際は、物理的な条件と心理的な機能によって取得する情報量が階層化されている。

まず物理的な条件として、空間的な位置関係に応じて情報量が決まる。自分からの一定の距離と方向に存在する他者の姿を見たり、声を聞いたりすることができ、他者との距離を縮めるとそれに比例して取得する情報量が増加する。

次に心理的な機能として、注目度に応じて情報量が決まる。人間の意識には志向性があり、無意識または意識的に何かを選択し、それについて多くの情報を得ている。選択される対象はその時点で関心を持っているものであり、誰かとコミュニケーションしている場合には、その相手に注目し、表情、視線、身振り、手振りなど細かな情報を得ている。

以上を考慮して、次の基本方針でアウェアネス情報の通信を制御する。(1) 現実世界での物理的な条件を模擬し、他者の存在と位置を認知するために必要な情報を仮想空間内での位置関係に応じてユーザに提示する。(2) 心理的な機能を模擬し、コミュニケーショングループに属するメンバーに関する情報は強調してユーザに提示する。コミュニケーショングループはオーラ方式を用いて近距離にいるユーザ間で形成する。図4では、A、B、Cはオーラが交差しコミュニケーショングループを形成しているのでインタラクションが強調される。

アバタ、音声、映像の各メディアがその特性により仮想空間においてのどのような働きをするかを検討し、ユーザへの提示の仕方を(1) - (2)の方針に基づいて制御する手法を以下に述べる。

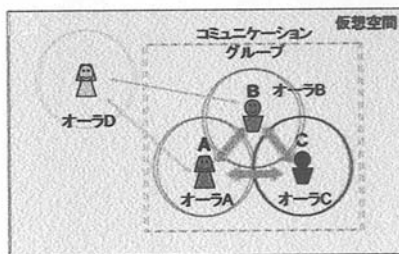


図 4: コミュニケーショングループ

3.3 仮想空間共有機能

遠隔地にいるユーザが、同じ場所に居合わせているるように位置関係を調整しながらコミュニケーションすることを実現するには、コンピュータとネットワーク上で仮想的に空間を共有する仕組みが必要である。

人間の認知過程では視覚情報と聴覚情報が大きな比重を占めているので、視覚的および聴覚的に空間情報を共有すれば、物理的には別々の空間にいても、同じ空間にいるかのように感じることができる。また人間は三次元の空間性によって現実感が生じるので、三次元の視空間と聴空間として仮想空間をユーザに提示すれば、その効果は高まる。仮想空間共有機能はユーザがコミュニケーションする場として、三次元の視覚空間を提供する。ユーザの視点(位置と方向)に基づいて、仮想空間を3Dグラフィックでディスプレイに表示する。ユーザはキーボードの矢印キーを押すことで仮想空間の中を自由に移動することができ、その動きに合わせてグラフィック表示も遷移する。

次に仮想空間の中でのユーザの存在を可視化し、ユーザの分身としての役割を果たすグラフィックであるアバタを配置する。アバタは仮想空間の中で他者と出会うことを可能にする。

アバタは、仮想空間にユーザが存在することを実体化したものであり、仮想空間におけるユーザの存在と位置を他のユーザに伝えることができる。また3Dグラフィックの表現によりある程度のリリアリティを持たせて表情や動作を伝えることができるが、ユーザの感情などを細やかに伝える手段としては不十分である。アバタは存在と位置のアウェアネスを伝えるものと捉えるのが適切である。

現実世界において他者の姿が見えるかどうかは、見る側の視野の範囲内に他者が存在しているかどうかで決まる。そこで仮想空間においては、ユーザの可視範囲(視野)としてフォーカスを設定し、フォーカス内に存在するアバタのみを見ることができるようとする。

3.4 音声通信機能

音声通信機能は、仮想空間の中で出会った他者と音声による会話を行うことができる機能である。現実世界に

おいて音声は、聞かれる側（音源）が発する音声の音量に基づいてその被可聴範囲が決まり、その中に他者が存在すると、その他者は音声を聞くことができる。そこで仮想空間においても音源の音量に基づいてニンバスを動的に設定し、ニンバスの範囲内にアバタが存在すると、そのアバタを操作するユーザは音声を聞くことができるものとする。これにより存在と位置を伝えるアウェアネスとして機能させる。

しかし音声の主要な役割は、言語によって会話するための媒体であり、また声色や抑揚によって感情や意図を伝えるアウェアネスである。そこでコミュニケーショングループに属するユーザの音声はより鮮明に聞こえるように制御し、現に会話している相手についてより多くの情報を取得できるようにする。

図5で、AはCが発する音声のニンバス内にいるのでその声を聞くことができる。またAとBはオーラが交差しているのでコミュニケーショングループを形成しており、互いの声を鮮明に聞くことができる。

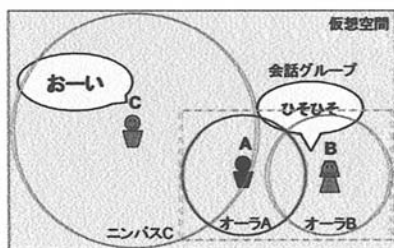


図 5: 音声通信の制御モデル

3.5 映像通信機能

映像通信機能は、Webカメラでキャプチャしたユーザの写実映像を表示することにより仮想空間で出会った他者の現実世界における表情や手振りなどの情報を認知することができる機能である。現実世界で表情や手振りに注目するのは、コミュニケーションしている相手に対してである。そこでコミュニケーショングループのメンバーになると映像を表示するように制御する。

以上のように、空間モデルに基づいて、メディアの特性に応じてユーザへの提示の仕方を制御することにより、現実世界に近い自然なコミュニケーションを実現できると考える。

4 フレームワークの実現方法

本章では、提案するフレームワークを実現するためのアーキテクチャについて設計の基本方針を述べる。

4.1 アーキテクチャの概要

アーキテクチャは、(1) 位置情報管理部、(2) GUI部、(3) 音声制御部、(4) 映像制御部により構成される。位置

情報管理部は仮想空間に参加しているユーザの位置情報を管理する。GUI部は、ユーザに対する仮想空間の表示とユーザの操作によるアバタの移動を管理する。音声制御部はユーザ間での音声会話を制御する。映像制御部はコミュニケーショングループに属するユーザ間での映像通信を制御する。各機構は位置情報管理部を中心に図6のように通信することで、それぞれの機能を実現する。各モジュールは、マルチプラットフォーム環境での利用を考慮してJavaで実装する。

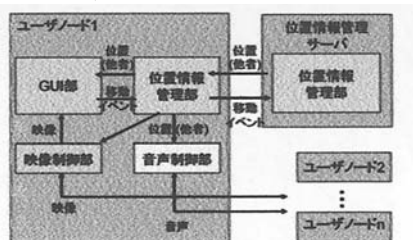


図 6: アーキテクチャの概要

4.2 位置情報管理部

位置情報管理部は、仮想空間におけるアバタの位置情報をユーザノード間で共有する仕組みを提供する。

位置情報の管理は、専用の管理サーバを設置し、クライアント/サーバ方式で行う。ユーザノードの位置情報管理モジュールは、GUI部がアバタ移動イベントを検知すると、それを管理サーバに通知する。管理サーバはクライアントからのイベント通知を基にアバタの位置を集中管理する。サーバは全てのアバタの位置情報をユーザノードに定期的に一斉同報で配布し、ユーザノードが他のアバタの位置を保持することを可能とする。

位置情報管理部は各アバタに対応する物理アドレス（ユーザノードのIPアドレス）も保持し、音声制御部や映像制御部がユーザノード間でユニキャスト接続の際に提供する。

ユーザノードの位置管理モジュールは、コミュニケーショングループの管理も行う。位置情報に基づいてオーラの範囲内にいるアバタを特定しグループに追加する。

4.3 GUI部

GUI部はユーザへの出力とユーザからの入力を制御する。

出力部は、仮想空間とアバタを3Dグラフィックでディスプレイに表示する。仮想空間の背景、オブジェクト、アバタの3Dデータはユーザノードにあらかじめ配布しておき、仮想空間におけるユーザの視点（位置と方向）と、位置情報管理部が保持する他のアバタの位置情報に基づいてそれらを描画する。3Dグラフィックの描画はゲーム開発用3DグラフィックライブラリであるJOGL (Java

bindings for OpenGL) を使用して実装する。

入力部は、ユーザの操作 (矢印キーの入力) によるアバタの移動を検知して、位置情報管理部に通知する。また出力部に対して新しい位置情報に基づく仮想空間画面の再描画を指示する。

4.4 音声制御部

音声制御部はユーザ間で音声による会話を行う際の、マイクからの音声の検出、ユーザノード間での音声データの通信、アバタの位置関係に基づく音響効果、スピーカまたはヘッドホンへの音声出力を制御する。

送信プロセスでは、マイクから検出した音声の音量と位置情報管理部が保持する他者のアバタの位置情報に基づいてニンバス内にあるアバタを特定する。次に特定されたアバタに対応するユーザノードにユニキャスト接続し音声データを送信する。

受信プロセスでは、受信した音声データに対応するアバタの位置情報に基づいて3D音響効果を付加し、スピーカまたはヘッドホンに出力する。コミュニケーショングループに属するユーザの音声である場合にはより鮮明に聞こえるように加工しカクテルパーティ効果を付加する。3D音響効果は、ゲーム開発用3DサウンドライブラリであるJOAL (Java bindings for OpenAL) を使用して実装する。

4.5 映像制御部

映像制御部は、Webカメラからのユーザ映像のキャプチャとコミュニケーショングループに属するユーザノード間での映像の通信を制御する。

送信プロセスでは、位置情報管理部からコミュニケーショングループへのメンバー加入の通知を受けて、対応するユーザノードにユニキャスト接続し、映像データを送信する。映像データはJMF (Java Media Framework) を使用してWebカメラからキャプチャする。

受信プロセスでは、受信したデータをバッファリングし、GUI部がこれを使用して図3のように仮想空間画面に表示する。

5 まとめ

本稿では、仮想空間においてアウェアネス情報を取り込んだコミュニケーションを実現するためのフレームワークを提案した。CSCWやグループ学習においては位置関係を調整することによる柔軟なグループ編成と、アウェアネス情報を伝達して人間味のあるコミュニケーションを実現することが必須である。アバタ、音声、映像のコミュニケーション媒体としての特性を活かして、オーラ、フォーカス、ニンバス等の空間モデルに基づいて通信を制御することにより、現実世界に近い自然なコミュニケーションを実現するための手法を提案した。

現在、アーキテクチャの設計方針に基づいてプロトタイプを実装している。音声制御部はRAVITAS[5]で実現されているのでこれを使用している。

今後は評価実験を行う予定である。評価の目的は仮想空間を共有し、アウェアネス情報を含むコミュニケーションを行うことで、コミュニケーションの効率・満足度がどの程度向上するかを調査することとし、TV会議との比較による主観評価を実施する。被験者グループに対してブレンストーミングなどの課題を与え、TV会議の場合と提案するフレームワークとで、どちらがよりコミュニケーションしやすいかを判定する。評価実験の結果を考察し、より自然なコミュニケーションを実現する上で効果的な制御モデルを検討していく。

謝辞

本研究の一部は、(財)中山隼雄科学技術文化財団の支援を受けて行った。ここに記して謝意を表す。

参考文献

- [1] Mehrabian, A.: "Silent messages, Implicit Communication of Emotions and Attitudes, 2nd Ed.," Wadsworth Pub. Co. (1981).
- [2] Hall, T. E.: "The Hidden Dimension Doubleday & Company," NY (1966).
- [3] Yang, Z., Cui, Y., Yu, B., Liang, J., Nshrested, K., Jung, S. and Bajscy, R.: "TEEVE: The Next Generation Architecture for Tele-Immersive Environments," Proc. of the 7th IEEE International Symposium on Multimedia (ISM'05) (2005).
- [4] Benford, S. and Fahlén, L. E.: "A Spatial Model of Interaction in Large Virtual Environments," Proc. of Third European Conference on CSCW (ECSCW'93), pp.107-122 (1993).
- [5] Yasumoto, K. and Nahrstedt, K.: "RAVITAS: Realistic Voice Chat Framework for Cooperative Virtual Spaces," Proc. of IEEE International Conference on Multimedia and Expo (ICME2005), (CD-ROM) (2005).
- [6] Dourish, P. and Bly, S.: "Porthole: Supporting Awareness in a Distributed Work Group," Proc. of Conference on Human Factors in Computing Systems, pp.541-547 (1992).
- [7] Gutwin, C. and Greenberg, S.: "A Descriptive Framework of Workspace Awareness for Real-Time Groupware," Computer Supported Cooperative Work, Vol.11, No.3-4, pp.411-446 (2002).
- [8] Gaver, W. W.: "The affordances of media spaces for collaboration," Proc. of ACM conference on Computer-supported cooperative work (CSCW92), pp.17-24 (1992).
- [9] Slavin, E. R.: "Cooperative Learning: Theory, Research, and Practice," Allyn and Bacon (1995).