人間と機械の「感性」コミュニケーション

対談 [ロボットやコンピュータになぜ「感性」を求めるのか]

菅野重樹 ― 土佐尚子

[すがの・しげき]「人間に役立つロボット」と「心を持つロボット」の両面から人間共存型ロボットの開発に打ち込む．早稲田大学理工学部機械工学科教授（53ページに略歴）

人間と共存し，役立つロボットとは

菅野●私は，人間と機械のコミュニケーションという視点からロボットを研究しているのですが，いくつかのアプローチの方法があって，その一つが人間と共存して人間の役に立つロボットをつくることです．この「ウェンディ」がそれで，人間のすぐそばで働くことができる，一緒に作業ができる．そういうロボットです．そのために必要なのは，まず絶対に安全であること，そして人間の動きに合わせることができて，人間とうまくコミュニケーションがとれることです．そういうロボットは今までなかったんですね．

土佐●この手なんか，すごくリアルですね．目もちゃんと二つついている…．

菅野●目で対象を認識して手で卵を割ることもできます．自分で動いて，床に落ちているものを拾うこともできる．とくに料理というのは，いろいろな要素が詰まっていて面白いテーマなのですが，いずれ人間と一緒に料理をしたり，一人で目玉焼きをつくって，後片付けくらいまでやれるようにしたいと思っているんですが．

土佐●移動は車輪ですね．二足歩行はすごく難しい…．

菅野●ええ．そこそこうまく歩けたとしても，倒れたときに人間に危害を与えかねない．コストも含めて，車輪で移動させた方がリスクが少ないですね．福祉や医療という目的を考えた場合も，バリアフリーが趨勢ですから車輪でいいのではないかと．

土佐●会話はできるんですか．感性的なコミュニケーションということに関わりますが．

菅野●早稲田大学のヒューマノイドプロジェクトでは，音声認識や音声合成の研究も進んでいるので，やがては音声的なコミュニケーション機能もつけられると思います．そこでその感性なんですが，人間とロボットが何か一緒に作業をしているときに，ロボットがちょっと目を動かしたり，首をかしげたりすると，すごくリアルな存在感が出てくるんですね．ロボットが感情を表わしているわけではないのに，人間がそこに感情移入するというか…．じゃあ，ロボットがどういうジェスチャーをすれば人間はどう感ずるのか，それは理論的にはよく分からない．研究テーマの一つです．

ロボットに「心」は発生するか

土佐●ロボット研究に人間の脳をシミュレートするものがありますね．

菅野●それに近いのがこっちの「ワメーバ」です．これは人間との心のコミュニケーション，つまり究極の人間共存ロボットを目ざすもので，その前提として，機械に心が発生するかどうかという壮大なテーマのもとに始まった研究です．そこで心とは何かと考えると，知・情・意の総体として捉えることができる．その根本にあるのが，脳幹など人間の脳の最も古い部分がつかさどっている情ではないか．これが今までのロボットには欠けていた．その情，つまり感情や情緒は，おそらく自己保存に関わるものです．たとえば私たちがうれしいと思うのは，生きていくためにいい刺激を受けたときの反応です．そういう情から知識を獲得したり，いろいろな行動を起こす意志も生まれてくる．そうしたプロセスを機械で実現しようというのが「ワメーバ」のプロジェクトなんです．

土佐●その心の動き，内部状態は何で表現するのですか．

菅野●普通のロボットのような外の環境を認識するだけでなく，回路の温度とか電流の状態とか，自分の内部状態をキャッチするセンサーがたくさんついています．その内部状態が自分が生きていくためにいいか悪いかを「自己保存評価関数」によって判断して，ホルモンを分泌する．もちろんコンピュータによるシミュレーションで，実際に何かの物質が流れるわけではないのですが，とにかく，そのホルモンの分泌によって，ディスプレイの色が変わったり，ある記号を声に出したりするわけです．頭と胸にあるディスプレイが「ワメーバ」の内部状態を表わすモニターなんです．

土佐●赤くなると怒っているとか…．

菅野●ええ．真っ赤になったり，真っ白になったり…．ただ，どういう刺激を与えるとどういう反応を示すといったモデルは与えていない．つくり込まない，というのが基本なんです．ですから，モニターに表示された色や声に出した記号を調べて，たとえば人間が小突くとどうもこの記号を出すようだ，これは怒っていることを表わしているんだな，とこちらが意味づけしていくわけです．

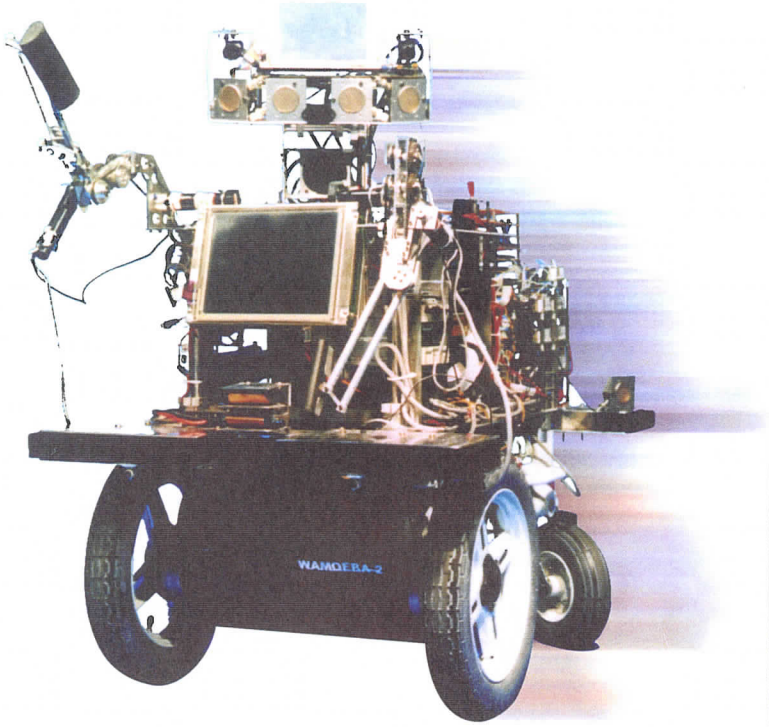
土佐●人間の感情と顔の表情との関係を分析し，データベース化して，それをロボットに載せる，そうした研究がありますね．それとは違う…．

菅野●それをやっているのはむしろ「ウェンディ」なんです．「ワメーバ」では，そういうモデル化によって疑似的な動きや表現をさせるのではなく，情緒とか感情の表現にしても，あくまでも「ワメーバ」に任せてある．つまりそれは，機械に心が発生するかどうかというテーマを原理的に追求したいからです．もちろん，ロボットに心を求めるなどナンセンスだという立場をとる研究者もたくさんいます．しかし私は，本当にロボットが人間と共存して動くためには，情が基本にあって，そのうえで知識や身体性を獲得していくような構造にしないとダメだと思うんですね．

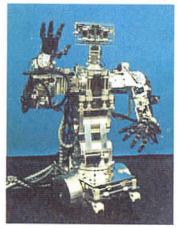
じゃあ，ロボットの情緒や感情とは一体何かということですが，生きているという構造だけを与えておいて，それが経験的に表現するようになったもの，それをわれわれが確かに感情を表わしていると感ずれば，それはロボットに感情があると見ていいのではないか．たとえば充電してやると活発に動き出す，「そんなにうれしいのか」と言葉を何度も投げ掛けてやると，ディスプレイの色が変わったり，ある記号を発生したりする．それは「うれしい」というメッセージではないか．人間だって，調子が悪いと，それを大脳がキャッチして不機嫌な表情をつくったり，何かそれを調整する機能を働かせたりするわけですから，そうした意味では，私は人間機械論者なんですね．

土佐●身体や人間の脳のいろいろな部分のガジェットをつくって，それを合成した方が早いような気もしますが…．

菅野●部分的にはそうなっています．でも，できるだけ自分で獲得させたいと思って頑張っているのですが，実をいうと「ワメーバ」の動きを評価しようとすると，われわれ自身，どこまでわれわれがつくり込んだもので，どこまでが自然発生的なものなのか区別がつかなくなってしまう．しかも，学習したり行動を起こすにしても，人間は何も命令しないので，「ワメーバ」自身がモチベーションを獲得しなければならない．そのための経験を積む機会，つまり身体機能や行動範囲が重要になるのですが，まだきわめて限られています．2歳の赤ちゃんくらいには育てたいのですが，育てるための枠組もまだ完成されていません．となると，一体いつになったら人間の役に立つロボットになるか分からない．ですから，やはり「ワメーバ」だけでは困るので，「ウェンディ」のプロジェクトも欠かせないわけです．



「ワメーバ」（左）
「ウェンディ」が「動作」に着目しているのに対し，ロボットに「心」を創出させる試みを行っているのが「ワメーバ」プロジェクト．知・情・意といわれる心の「情」の部分に着目し，それを自己保存に関わるものと位置づけ，快，不快という自己保存に関係する心の状態＝感情を，あらかじめ埋め込まれた自律系の状態をモニターで示すことによって表現する．1996年に開発された「ワメーバ2」では，モーター出力調整，カメラ視野角度調整，体内回路の温度調整などの機能が，人間のホメオタシスに相当する自律系として採用されている．センサーは外部環境と内部状態をセンシングするものを併せて11カ所．自律系の状態を評価する「自己保存評価関数」は7種類．頭部と胸部の液晶パネルによって，さまざまな「感情」を表現する．「ワメーバ」の名はアメーバに早稲田のWを冠したもの．



「ウェンディ」（上）
早稲田大学のヒューマノイドプロジェクトの一環として，「人間の身近にいて人間に役立つロボットをできるだけ早くつくること」を目的に開発されたのが「ウェンディ」である．人間との相互コミュニケーションを図るために，指先や目の動きなどの動作をできるだけ人間に近づけることを目指している．基本的な動作基準は人間がプログラミングして組み込み，ロボットはセンサーから得た情報を自身で判断し，制御パラメータに従って動く．指先はかなり器用で，卵を割ったりタバコをはさんだりすることができる．身長150cm，体重約100kg．車輪で移動し，胴体を衝撃吸収材で覆うなど，人間の安全を考えて設計されている．早稲田のWを頭文字にした「ウェンディ」は，その名の通り「女性」である．

コンピュータと人間とのインタラクティビティの追求

土佐●「ワメーバ」では，人間の側が「つくり込まない」とおっしゃいましたが，すごく分かるような気がします．私は，コンピュータやサイバースペースと人間との感性的なレヴェルでのコミュニケーションをテーマに研究や作品づくりをしていますが，作品づくりでいつも考えているのは，どうやって人間に感情移入させるかということなんです．そのためには，やはりあまりつくり込んでしまってはダメなんですね．人間とのギャップばかり目立って陳腐なものになってしまう．であれば，いかにしてコンピュータというシステムによってつくられたオブジェクトやスペースに人を引き込んでいくか，そこがすごく難しいんですね．

菅野●感性をモデル化して，それをつくり込んでしまうとウソっぽくなっちゃう．何かこう，ひょっと出てくるものではないと…．

土佐●コミュニケーションというのは，つまるところインタラクションですからね．「ニューロベイビー」というキャラクターから始まって，だんだんシステムが大きくなり，映画や演劇の手法を使った「インタラクティブシアター」までつくってしまいましたが，どんな作品であっても，テーマはインタラクティビティをどう実現するかなんです．

菅野●「インタラクティブシアター」というのはどういものなんですか．

土佐●その一つが国際ベルリン映画祭のニューメディア部門で上演した「ロミオ&ジュリエット黄泉にて」という作品で，簡単にいうと，磁気センサーを着た観客がヒーローとヒロインを演じ，台詞を自分で選択して，感情を込めてそれを読んだり，いろいろなアクションを起こすことによって，サイバースペースのキャラクターが感情と音声，動作に呼応し，リアルタイムにストーリーを生成させていくというものです．画面上にはロミオとジュリエットのアバター（化身）がいて，演ずる人の感情表現によってアバターの表情も変わる．アバターの動きは磁気センサーウェアによって制御される．動作音も仕組まれているので，演ずる人はアバターがまるで自分であるかのように感ずるようになる．そうなるとしめたもので，サイバースペースに引き込まれていって，キャラクターとの感性的なコミュニケーションも没入感とリアリティが生じます．

菅野●これまでやられていたことの集大成という感じですね．

土佐●ええ．「インタラクティブポエム」では，コンピュータに言語的な対話認識機能と，　人間の声が持つ感情の認識機能を持たせ，連歌型の詩を通して，人間とコンピュータとのインタラクティブなコミュニケーションの可能性を追求したわけですが，「インタラクティブシアター」は動き

[とさ・なおこ]インタラクティビティをテーマとした研究と作品制作に取り組む．ATR視聴覚機構研究所客員研究員（54ページに略歴）

「ニューロベイビー」1993年〜'95年
人の話しかける声から感情を抽出し，インタラクティブに答えるひとがた感情キャラクター．人間とコンピュータの感情によるコミュニケーションに着目し，インタラクティビティはコミュニケーションであるとする立場から，インタラクティビティを持った感情表現のモデル化を試みた作品である．技術的には音声処理技術とニューラルネットワーク技術による話者の感情認識とそれに反応する顔の表情合成からなる．

「インタラクティブボエム」
1997年〜'98年
感情を中心としたノンバーバルインタフェースと言語的なバーバルインタフェースの融合を目ざして開発された「インタラクティブボエム」は，声からの感情認識機能に言語的な対話認識機能を付け加えることによって，コンピュータとの間の声による感情のやりとりだけではなく，言葉の意味も同時に理解して連歌形式で人間と即興詩をつくれるコンピュータ詩人システムである．詩を朗読しあうことによって互いの感情移入を図ることを特徴としており，対話型の新しいインタフェースデザインとして注目される．

や表情も入っているし，キャラクターも複数で関係性がすごく入り組んでいる．私のやろうとしているのは，芸術の手法を使った工学なんですが，映画でも演劇でも，素晴らしいものは完璧に感情移入させますよね．それをコンピュータのシステムの中で実現しようという試みです．つまり，表現になっていて，かつ数理的なシステマティックな構造をもっているわけです．それはすごく大変で，トライ・アンド・エラーの連続でした．とくに難しいのは，たとえば戦闘シーンとか，やるかやられるかでストーリーが変わる．そういうダイナミクスをどのように設定するかということです．

無意識のコミュニケーション

菅野●いや難しいと思います．言語や音声，動きといった認識機能をセットして，それを枝分かれしていくシナリオに当てはめて，しかも映像ともつなげるわけでしょう．その構造って，とんでもなく複雑になるんじゃないですか．頭の中で整理して考えられるんですか?

土佐●映画制作のように，自分である程度ディレクションして，それからいろんな専門分野の人たちと一緒につくっていきます．それでも細かいところまではとても制御できないんです．まず時間的なタイミングを合わせることが難しい．アナログの映画にはタイムコードがあって，それで俳優の動きやBGMなど，すべてを合わせていきますね．インタラクティブシネマの場合は，結局フレームの容量，つまりコンピュータの速度に規定されてしまうわけです．

菅野●それが人間のリズムに合ったり合わなかったり…．

土佐●そう，ノリ．

菅野●それはロボットの世界でも注目されているテーマなんです．一方が一方に合わせるのではなく，二人の歩調が自然に合うように，何となく双方が歩みよって落ち着くようなあるダイナミクスにおけるリズムがある．それを「引き込み現象」というのですが，ロボットと人間のインタラクティブなコミュニケーションを考えた場合，これはすごく重要なことで，映画みたいな一方向だとインタラクションにならないんです．

土佐●それで次は，キャラクターやアバターをマネキンのような立体にして，三次元の映像と合成してみたいんです．インタラクションがもっとリアリティを持ってくるのではないかと．今までは人間がサイバースペースに行っていたけれど，反対に現実空間に仮想物体が入ってきて人間とコミュニケーションする．ロボットにちょっと似てますが…．

菅野●それは面白い．体感が加わりますからね．バーチャル・リアリティでも触覚や体感的な刺激がまだうまく取り込めないんですが，インタラクティブなコミュニケーションにとって体感というのはすごく重要な要素なんですね．

土佐●それと人間って嘘をつきますよね．だから言語とか音声で感情をセンシングして，それでコミュニケーションするというのは限界があると思うんです．本当に感情とか感性を捉えるには無意識のレヴェルに着目すべきでないか．そう思ってつくり始めたのが「無意識の流れ」という作品です．これはノンバーバルコミュニケーションの様子を画像化したもので，男女二人の観客の気持ちを心拍数と手の動きからセンシングして，たとえばニッコリ笑っていてもじつは互いにあまり関心がないとか，その反対であるとか，そうした隠された心理や感情を，二人の分身であるCGの人魚によって表現しようという作品なんです．

菅野●言葉はコンテキストに依存しているので，必ずしもその人の感情を反映しない．その点，生理情報というのは非常に正直ですね．私たちの研究でも，人間に合わせられる機械を考えた場合，その人間というのは人間一般ではなくて個人なんです．そうなるとどうしても生理情報が必要になる．それで人間にセンサーをつけて情報をキャッチし，たとえばもっと人間がリラックスできるように機械の側のパラメータを変えてやる．そういう研究を実際やっているわけです．

土佐●私の作品でも，初期の音声認識によるコミュニケーションとかは，人間の側が無理をして，コンピュータに合わせている面があるんです．本当はコンピュータがもっと人間の感情を理解すべきなんですけどね．

菅野●ただそれを固定されたシステムとして組むことは可能ですが，ロボットの場合は非常に難しい．個体として動くので人間の生理情報の取りようがないからです．人間にベタベタとセンサーを貼りつけるわけにもいかない．だから今のところは，ロボットが自分の状態を何となく表現して，人間がそれに合わせている．ロボットはすごくわがままなんです．

なぜ機械に感性を求めるのか

土佐●早稲田のプロジェクトの中に，指揮者に合わせてマンボを踊るロボットがありますね．あれなどは，指揮者の気持ちを理解しているというか，いかにも感性をもって踊っているように見えますね．

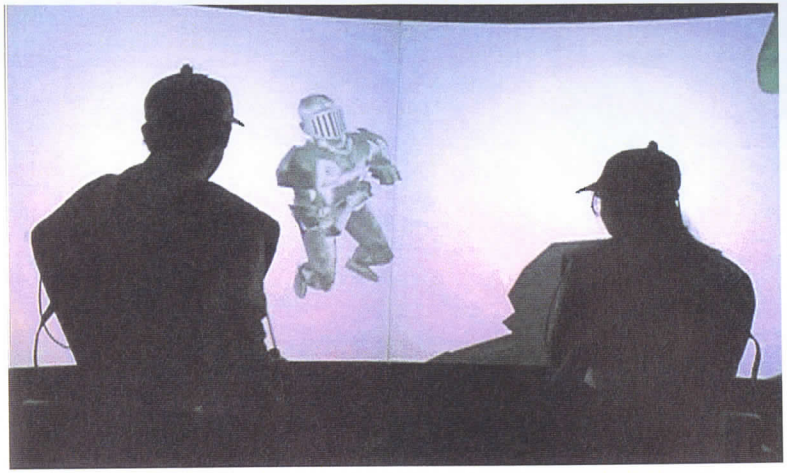
菅野●確かにそう見えるかも知れませんが，そんなに複雑なものではないんです．果たして感性をもっているといえるかどうか…．

土佐●そこですごく大事だと思うのは，コンピュータの得意なものは何かということです．それも，よくいわれるような，コンピュータがやっていることは結局は計算に過ぎないのだからとか，そういう薄っぺらな議論ではなくて，感性といったものに入り込んでも，コンピュータの得意なことはあると思うんですね．人間の感性にはばらつきがあるし，言葉が違えば，その何となくの感性でしかコミュニケーションできない．たとえばそこに感性の閾値，つまり何かしらの基準のようなものがあって，感性翻訳のようなものができるようになれば，人間のコミュニケーションにとってすごく有効だと思うんです．ひょっしたらそれは，コンピュータが得意とすることかもしれない．いろいろな文化の違いもデータベース化されていて翻訳してくれる．そうした可能性も決してないとはいえないと思うんですね．

菅野●エンジニアリングの観点から考えると非常に難しいですね．どうしてもそのモデルを人間が記述しなければならない．おそらくそれは不可能でしょう．となるとコンピュータやロボットに獲得させるしかないわけですが，今度はその手続きをどう数値化して入れるか，それが難しい．堂々巡りになっちゃいますね．

土佐●方法論を見出すまでが大変ですね．

菅野●私の研究室で感性をダイレクトに扱った研究もあります．その一つがピアノ演奏ロボットで，これは1985年の科学万博にも出品されて毎日演奏していたのですが，ほぼ完成の域に達しています．そこで次はヴァイオリンだということになったのですが，これがすごく難しい．ヴァイ



オリンは弓の速度や弦を押し付ける力，弓と弦の接する位置などによって音量も音色も変わる．人間が弾くのも非常に難しくて，運動だけ見れば機械の方が得意だともいえます．ところが人間の身体機能的な制約が，温かい音とか冷たい音，すごく豊かな音といった，感性的な表現を生んでいるともいえるわけです．そこでまずは，感性的な表現と運動のパラメーターを結びつけて演奏させたわけですが，これだけではまだ音楽になりません．おそらく人間の演奏には微妙な揺れとか，そういうものがあって，それが何か人間的なフィーリング，独特の表現を生んでいる．それが機械にはなかなかできないんですね．

土佐●正確に弾きすぎちゃう…．

菅野●そう．だから「ディープブルー」はチェスのチャンピオンになったけれど，ロボットはまだとてもヴァイオリニストにはなれない．チェスに勝つという目的だけを考えれば，とりあえず感性は必要ないんですが，ヴァイオリンの演奏はそうはいかない．

土佐●でも，これから人と機械との接触が日常的にますます多くなることは確かです．となると，コミュニケーションということを考えた場合，コンピュータやロボットに感性的なものを持たせることは絶対に必要だと思います．

人間，そしてもう一つの存在…

菅野●先ほど文化の違いとおっしゃいましたが，面白いというか，難しいというべきか，ロボットの場合は，それに対する捉え方そのものが文化によって全く違うんです．日本人は非常に人間に近い存在として感情移入する．アミニズムの世界に入ってしまうんですね．だから地雷を探知させるとか，そういう話になると，ロボットにそんな危険なことをやらせるの，なんてことにもなりかねない．それに対して欧米人は，基本的に機械である，道具である，と考えていると思います．人間と対話し，共存するようなロボットの研究というのは，実は日本が一番盛んなんです．

土佐●ペットロボットが発売されましたが，フランス人の研究者が言っていました．日本人はこういうの好きだね，フランス人はこんなこと絶対に考えないって．

菅野●たとえばMITにもヒューマノイドロボットのプロジェクトがありますが，その目的はあくまで人間の知性が生まれるメカニズムを知ることです．知性が外界とのインタラクションで形成されるとしたら，コンピュータに身体をもたせないと分からない．そういう考えであって，決して人間の仲間をつくろうとしているわけではないんです．

土佐●私たちのやっていることも，結局は人間をどう考えるかということだと思うんです．ただ，一番の関心事はコミュニケーションなんですね．人間同士のコミュニケーションは謎の部分です．ある言葉を交わしていても，お互いが同じビジョンを共有しているかどうか分からない．それがある瞬間バーンとクロスしたり…．そういう関係は，コンピュータとコミュニケーションしているとすごくよく見えてくる．ですから，コンピュータに人間らしさを与えて，それとのコミュニケーションを研究することは，人間同士のコミュニケーションの問題に必ずつながっていくと思うんです．

菅野●ええ．人間の役に立つロボットをつくると言って，そこに感性や人間とのインタラクティビティを求めるのは，やはり人間そのものを考えたいからなんですね．

土佐●私がなぜコンピュータを使ったキャラクターとか，そういう作品をつくろうとしたのかを考えてみると，自分の分身というか，自分と他人の中間にいるような存在，そういうものが欲しいと思ったことがきっかけなんです．人とのコミュニケーションって結構エネルギーがいるでしょう．相手もエネルギーを発散してくるし，元気じゃないとできない．でもいつも元気で快活で，というわけにはいかない．そんなときに話をしたりサジェッションしてくれるような相手が欲しかったんです．先生がロボットの研究を始めたきっかけは?

菅野●さて何でしょうね．高齢化社会ですから，私たちが年をとったとき，一緒に遊んでくれてたり，話し相手になってくれるロボットがいると安心かな，と．これはまあ福祉ロボットの発想で，ちょっと建前かな．むしろ自分と違うんだけど自分に属するような相手，そういものをつくりたい…．

土佐●私と似てませんか．

菅野●そういわれれば…，うん，そうかな?

感情を認識・生成するインタラクティブシアター
「ロミオ＆ジュリエット黄泉にて」
1998年〜
物語性をもつコンピュータドラマと観客との感情移入型インタラクションの融合を目ざした「インタラクティブシアター」は，感情に反応する擬人化キャラクターが物語に登場し，観客自身が主人公を演じることによって没入感と感情移入が図られ，またドラマの進行そのものにも観客が参加できるインタラクティブストーリーの導入が試みられている．この概念に基づくコンテンツ例として制作された「ロミオ＆ジュリエット黄泉にて」は，1999年度国際ベルリン映画祭「ニューメディア部門」に選出され，上演を行なった．

「無意識の流れ」1999年〜
男女の人魚CGが二人の観客の分身（Agent）として働く．分身の人魚は，観客の頭骨に装着している電極から心拍数に同期して動いており，心拍数をPCで計算したシンクロ度インタラクションモデルによって，人魚たちは二人の観客相互の隠された次元のノンバーバルコミュニケーションの様子を表現することになる．例えば，お互い関心があるのだが，はずかしいので，無関心を装っている．しかし，心臓はドキドキしているなど…．さらに二人の手の動きを設置されたカメラがとらえ，PCで画像解析を行ない，シンクロ度インタラクションモデルに応じて，シンクロ度の高い場合は体験者の手の動きにお互いの分身のCGが追従し，シンクロ度の低い場合は分身のCGが逃げていく．また，お互いの分身の人魚に触れた場合，バイブレータで疑似的に触感を感じることができる．