

Spline 上のオーディオ・レンダラの実現

4 F - 4

前田 慎司 小塚 宏 福岡 久雄 下間 芳樹

三菱電機株式会社 情報技術総合研究所

1 はじめに

Spline[1] は、遠隔地に分散した複数のユーザがネットワークを介して仮想世界を共有する分散仮想環境構築基盤ソフトウェアである。Spline のアプリケーションは、Spline が提供する API を用いてオブジェクトを生成し、仮想世界の部分世界を構築することができる。Spline はマルチキャスト通信を用いて仮想世界の情報(世界モデル)を同一セッションに参加しているアプリケーション・プロセスに複製し、緩やかな一貫性の維持を行い、仮想世界の共有を容易にしている。

我々は、分散仮想環境において、ユーザ同士が音声により対話をしたり、仮想世界内に配置されたサウンドをユーザに提供することを可能とするオーディオ・レンダラを Spline 上に開発した。本論文では、その実現方法と概要について述べる。

2 Spline のオーディオ機能

Spline を利用した分散仮想環境では、ユーザは仮想世界にアバタとして入り込み、アバタを操作することにより仮想世界内を自由に動き回ることができる。Spline は、仮想世界をより豊かに表現するために、以下のようなオーディオ関連のオブジェクトを提供しており、世界モデルで管理している(図 1)。

- spAudioSource
音源の属性を持つオブジェクトである。位置の指定により仮想世界内に配置され、音を発生させる。
- spHearing
耳に相当するオブジェクトである。位置情報を持ち、仮想世界内の音源(spAudioSource)が発生させる音の観測点となる。アバタがspHearingを持つ場合にのみ、ユーザは音を聞くことができるようになるため、必要に応じてサウンド機能を On/Off することが可能である。
- spSpeaking
口に相当するオブジェクトである。位置情報を持ち、ユーザの発話による音声の発生点となる。アバタがspSpeakingを持つ場合にのみ、ユーザは発話が可能となる。

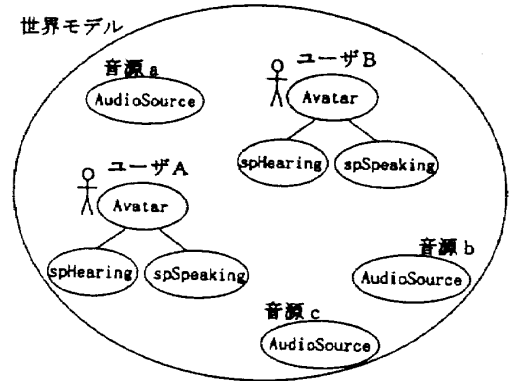


図 1: 世界モデル内の音源とアバタのオブジェクト

3 オーディオ・レンダラの実現

Spline のオーディオ・レンダラは、世界モデル内のオーディオ関連のオブジェクトの情報を基に、仮想世界の様々なサウンドをユーザに提供する Spline アプリケーションである。

あるイベントに対する効果音や背景音のように、予め録音可能なサウンドデータを仮想世界に配置する場合、サウンドデータを事前にユーザが使用する計算機に配布しておくことで、実行時のデータ転送を低減できる。つまり、音源とアバタの相対位置を基に音量・音質を調整し、複数の音源を合成して再生することにより実現することができる。

一方、ユーザの音声対話は、ある程度の実時間性を持たせながら、他のノードにネットワークを経由して、音声データを転送する必要がある。

従って、オーディオ・レンダラには、サウンドの音量・音質の調整、複数の音源の合成、ストリームデータのネットワーク転送の3つの機能が必要となる。

3.1 音量・音質の調整

オーディオ・レンダラは Spline アプリケーションであり、他のプロセスと同様にセッションに参加し、世界モデルの複製を持つ。このため、オーディオ・レンダラは世界モデル内の音源(spAudioSource)の情報、およびユーザが操作するアバタの情報を得ることができ、これらの情報を基にして各音源の音量を調整している。

各音源は内径と外径を設定することにより、音の広がり方を定義することができる。すなわち、ユーザの位置

が内径より内側にいる時に音量が最大、外径より外側で音量が最小(ゼロ)となり、内径と外径の間は滑らかに音量が変化する(図2)。

世界モデルは一定時間(例えば200msec)毎に更新されるため、アバタと各音源の相対距離を算出して音量を決定する処理は、この更新のタイミングで行われる。

また、現在は音質の調整は行っていない。

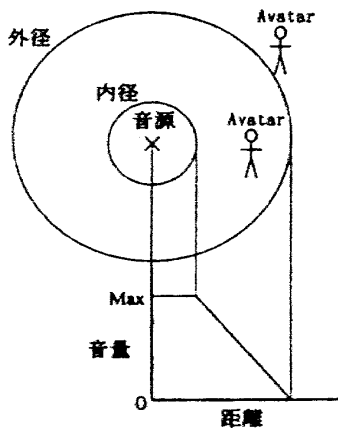


図2: 距離に応じた音量調整

アバタの口(spSpeaking)から発せられる音声データは、他の音源(spAudioSource)と同様に音量調整の処理を行っているため、ユーザ同士の音声対話についても、ユーザ間の距離が音量に反映される。

3.2 複数の音源の合成

Spline およびそのオーディオ・レンダラの開発は現在、WindowsNT4.0上で行っており、サウンドの出力にはMicrosoft DirectXに含まれるDirectSoundを使用している。このため、複数の音源の合成処理は、各音源毎にサウンドバッファにサウンドデータを複製することで自動的に行われる。

3.3 ストリームデータのネットワーク転送

ストリームデータ(今回の場合は音声データ)のネットワーク転送は、LBNLによって開発されたネットワーク音声会議ツールであるvat[2]を基にして、オーディオ・レンダラに組み込んでいる。vatは通信プロトコルとしてRTP(Realtime Transport Protocol)を使用しており、マルチキャスト通信を用いたネットワーク経由での音声対話が可能である。音声圧縮アルゴリズムとして、PCMの他にGSMやLPC等の低速回線に適したアルゴリズムも使用することができる。

オーディオ・レンダラに組み込むためにvatに追加したコードは、発話者をSplineの音源(spAudioSource)と同様に扱い、音声データを距離に応じて音量を調整して

再生する処理、および音声の出力にDirectSoundを使用するようにした点である。

4 評価

オーディオ・レンダラをSplineアプリケーションの1つである自転車メディアパーク[3]に適用し、2地点(鎌倉と横浜)をISDN回線(128kbps)で接続し、評価を行った。自転車メディアパークでは、約20個の音源がパークの各所に配置されている。ユーザは、水音、鳥のさえずり、ジャズ音楽等、場所に応じたサウンドを聞くことができ、またユーザ同士はお互いに音声対話を楽しむことができる。

結果として、ユーザが公園内を移動すると、各場所に配置されたサウンドが距離に応じて滑らかに音量が変化しながら再生され、また、2人のユーザがお互いに音声で対話することができた。

ただし、音声対話に関しては、音声の圧縮伸長をソフトウェアによって行っているため、0.5秒程度の遅延が生じていた。また、使用するネットワーク帯域は、音声の圧縮にGSMを使用した場合、参加ユーザ1人当たり約18kbps(パケット化のオーバーヘッドを含む)であった。Spline自身も世界モデルのデータ更新等にネットワークを使用しているが、ISDN回線を使った2ユーザでの音声対話は可能であることが確認された。

5 おわりに

我々は、分散仮想環境において、ユーザ同士が音声により対話したり、仮想世界内の場所に応じたサウンドをユーザに提供するオーディオ・レンダラをSpline上に開発した。

現在は1つのマルチキャストアドレスを用いて、ユーザ全員の音声データを流している。このため、ユーザ数に比例してネットワーク負荷が増加してしまう。今後は、公衆網を利用して多人数が参加できるように、複数のマルチキャストアドレスを動的に割り当て、お互いに対話しているユーザだけが同一のマルチキャストアドレスを使用するように改良することを考えている。

また、距離だけではなく方向の定位ができるように、よりリアルな3次元サウンドを提供すること、音声品質の改良、遅延時間の解消等もこれからの課題である。

参考文献

- [1] R.Waters 他: 分散仮想環境基盤ソフトウェア Spline, 三菱電機技報 (Feb 1997).
- [2] vat, <http://www-nrg.ee.lbl.gov/vat/>
- [3] 高橋他: Spline上の仮想公園 - 自転車メディアパーク -, 日本バーチャルリアリティ学会研究報告 Vol.2, No.1(Nov 1997).