

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Μαστοράκης Ιωάννης

Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτ. Εργασίας: Αναπλ. Καθηγητής Α. Μουχτάρης

Παρασκευή, 6/11/2015, 10:00

Αίθουσα Β108, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

**"Η υλοποίηση του ImmACS - ένα τηλεπικοινωνιακό σύστημα με υποστήριξη
εικονικού ήχου"**

ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή, υπολοποιούμε ένα ολοκληρωμένο τηλεπικοινωνιακό σύστημα το οποίο επιτρέπει τη καταγραφή και την αναπαραγωγή εικονικού ήχου υψηλής ευκρίνειας. Το σύστημά μας βασίζεται σε μια εύρωστη τεχνική με χαμηλό υπολογιστικό κόστος, που χρησιμοποιεί μία κυκλική συστοιχία μικροφώνων για την καταγραφή του ήχου, και ακουστικά ή ηχεία για την αναπαραγωγή. Το μοντέλο επικοινωνίας υποστηρίζει πολλαπλούς, ταυτόχρονα ενεργούς χρήστες, που μπορούν να στέλνουν και να λαμβάνουν εικονικό ήχο. Για την διανομή των ροών πληροφορίας χρειάζεται η μεσολάβηση ενός διακομιστή. Υποστηρίζεται επίσης προαιρετικά και η επικοινωνία μέσω βίντεο.

Υλοποιούμε το λογισμικό του χρήστη τροποποιώντας το Baresip, μια εφαρμογή ανοιχτού κώδικα για VoIP κλήσεις. Υλοποιούμε όλες τις λειτουργίες για τον εικονικό ήχο σε μια βιβλιοθήκη, την οποία σχεδιάσαμε για ευελιξία και μέγιστη απόδοση σε πραγματικό χρόνο. Ενσωματώνουμε την βιβλιοθήκη στο Baresip και τροποποιούμε την διεπαφή του ανάλογα. Επιπρόσθετα, αλλάζουμε το Baresip ώστε να μπορεί να λαμβάνει πολλαπλές ροές δεδομένων ανά κλήση, ώστε να είναι δυνατή η διαχείριση της ροής του κάθε χρήστη ξεχωριστά. Η ροή εικονικού ήχου αποτελείται από ένα κανάλι, συνοδευόμενο

από χωρικά δεδομένα, τα οποία εισάγουμε στο κάθε πακέτο ως επιπρόσθετη πληροφορία. Τέλος, υλοποιούμε την αποκωδικοποίηση των ροών παράλληλα, για να αυξήσουμε την απόδοση στα συστήματα πολλαπλών επεξεργαστών.

Ο διακομιστής βασίζεται επίσης στο Baresip. Ο κάθε χρήστης συνδέεται στο διακομιστή ως χρήστης-προς-χρήστη, εκμεταλευόμενος την ήδη υπάρχουσα υποστήριξη του Baresip για αυτό. Ωστόσο, απενεργοποιήσαμε την αποστολή των τοπικών ροών στο διακομιστή, και τον τροποποιήσαμε ώστε μόνο να μεταβιβάζει τα πακέτα μεταξύ των χρηστών που είναι συνδεδεμένοι στον ίδιο λογαριασμό. Με αυτό το τρόπο, υποστηρίζονται πολλαπλά δωμάτια τηλεδιάσκεψης, αφού κάθε λογαριασμός συμβολίζει και ένα διαφορετικό δωμάτιο. Τα πακέτα αναμεταδίδονται απείραχτα, χωρίς να υλοποιείται κάποια μετατροπή ή μίξη, καθιστώντας τη παρουσία του διακομιστή ανεπαίσθητη στην όλη επικοινωνία. Επιπρόσθετα, παραλληλοποιούμε την αναμετάδοση των πακέτων για να πετύχουμε μεγαλύτερη απόδοση στα συστήματα πολλαπλών επεξεργαστών.

Διευκολύνουμε την διαχείριση των ροών ήχου, προσφέροντας μια ευέλικτη γραφική διεπαφή. Η διεπαφή είναι μια αυτόνομη εφαρμογή και επικοινωνεί με τη κύρια εφαρμογή μέσω δικτύου. Επιτρέπει την παρακολούθηση και την προσαρμογή, σε πραγματικό χρόνο, της κατεύθυνσης και της έντασης των ηχητικών πηγών. Όλη η επεξεργασία σήματος πραγματοποιείται στη κύρια εφαρμογή, ενώ η διεπαφή προσφέρει μόνο μετα-δεδομένα που περιγράφουν τα ηχητικά φίλτρα.

Τέλος, για να απλοποιήσουμε τις δοκιμές του συστήματος, επιτρέπουμε την φόρτωση και την αναπαραγωγή ήχου από και προς αρχεία στο δίσκο. Έτσι μπορεί εύκολα να στηθεί ένα πλήρως ελεγχόμενο πειραματικό περιβάλλον, χωρίς να χρειάζεται κάποια εξωτερική συσκευή. Στη ίδια βάση, προσφέρουμε μια ειδική λειτουργία όπου ο διακομιστής αναμεταδίδει τα πακέτα πίσω στους αποστολείς τους.

Mastorakis Ioannis

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Associate Professor A. Mouchtaris

Friday, 6/11/2015, 10:00

Room B108, Computer Science dept., University of Crete

“The implementation of ImmACS - an Immersive Audio Communication System”

ABSTRACT

In this work, we realize a complete communication system that allows the capturing and reproduction of high-quality immersive audio. Our system is based in a computationally efficient yet robust technique, that utilizes a circular microphone array for audio capturing, and headphones or loudspeakers for audio reproduction. The communication model supports multiple, concurrently active clients, that can simultaneously transmit and receive immersive audio. The model requires the mediation of a server for the distribution of the streams. Optional video streaming is also supported.

We implement the client by modifying Baresip, an open-source VoIP client. We realize all the immersive audio functions in a library, which we have designed for optimal real-time performance and flexibility. We incorporate the library into Baresip and extend its configuration interface accordingly. Furthermore, we modify Baresip so that it is able to receive multiple streams per call, so as to allow the adjustment of the stream of each peer separately. The immersive audio stream consists of one audio channel accompanied by spatial meta-data, which we explicitly include in every packet as side information. Lastly, we utilize multiple threads for the decoding of the streams, so as to increase the throughput of the system in multi-core systems.

The server is also based on Baresip. Each client connects to the server as peer- to-peer, exploiting the fact that Baresip is by default a VoIP client. However, we have disabled the transmission of the local streams at the server; it only relays the packets between clients connected to the same SIP account. In this way, multiple conference rooms are supported, as each SIP account represents a different conference room. The packets are relayed completely untouched, with no transcoding or down mix taking place, making the server's presence transparent to the overall communication. In addition, we utilize several worker threads that perform the relaying of the packets, so as to take advantage of the multi-core systems.

To facilitate the manipulation of the audio streams at the client, we provide a flexible graphical interface. The interface is an autonomous application that uses network sockets to communicate with the client. It allows the monitoring and adjustment, in real-time, of the direction and volume of the audio sources. All signal processing is performed at the client, while the interface only sends meta-data describing the audio filters.

Finally, to accommodate for system testing, we provide the ability to capture and reproduce audio from and to files at disk. In this way, a fully controlled input and output environment can be set easily, without the need for external hardware. In the same basis, we provide a special "echo" mode of function, where the server transmits the packets back to their senders.