

Résumé

La transmission de la voix sur Internet (VoIP) se fait par émission et réception de paquets. Au récepteur, certains paquets manquent à cause des délais, de la congestion ou des erreurs de transfert. Cette perte de paquets dégrade la qualité de la voix au niveau du récepteur. Des algorithmes de dissimulation des pertes (concealing) sont utilisés au niveau de l'émetteur ou au niveau du récepteur afin de combler ces pertes.

Dans ce travail, nous avons mis au point et implémenté un codeur harmonique, véhiculant la voix sur IP. Ce codeur est conçu en utilisant la prédiction linéaire à excitation mixte (MELP).

Etant donné que la transmission de la voix est effectuée en temps réel, le récepteur ne peut pas requérir à la retransmission des paquets perdus à cause des délais de transfert trop importants. Nous avons alors utilisé la technique de codage par description multiple (MDC), qui est basée sur la redondance des informations. Ce sont ces redondances que l'on appelle descriptions. Ces dernières sont transportées dans chaque paquet à partir duquel les trames perdues seront reconstituées. Ainsi, nous avons codé le signal sur deux descriptions : la première permet de coder la trame courante à 2.4 kbps. Elle sert à la reconstruction du signal avec une bonne qualité. La seconde description utilise un autre codeur MELP, mais fonctionnant à un débit de 1.2 kbps. Ce dernier code trois trames successives dans le même paquet, à savoir les trois trames qui succèdent la trame courante. Cette description contribue à reconstruire la parole lorsqu'un, deux, voire trois paquets seront perdus. Nous pouvons ainsi récupérer jusqu'à trois trames lorsqu'on perd trois paquets successifs. L'objectif est de garder l'intelligibilité et une certaine qualité de la parole lorsque le taux de perte de paquets augmente.

Pour nos simulations, les pertes de paquets se font de façon aléatoire. A la réception, et dans le cas sans perte, on forme la voix à partir de la première description (codeur MELP à 2.4 kbps). En cas de perte, la seconde description (MELP à 1.2 kbps) reçue précédemment, et préalablement mémorisée au niveau du récepteur, servira à la reconstitution.

Pour valider notre méthode, nous avons utilisé un matériau linguistique formé de corpus multilingue. Le premier est composé de phrases arabes phonétiquement équilibrées conçues au niveau de notre laboratoire. Pour les langues française et anglaise, nous avons utilisé les phrases célèbres phonétiquement équilibrées : « la bise et le soleil » et « the sun and wind ». L'évaluation de la qualité de la parole utilise la méthode PESQ, normalisée ITU-T, conjointement à des tests d'écoute. Les résultats obtenus montrent que la méthode proposée améliore considérablement la qualité de la voix reconstruite lorsque celle-ci est véhiculée sur IP.