



HAL
open science

Couplage et exploitation d'un synthétiseur à formants sur un ordinateur

Philippe Abbeg, Robert Espesser, Bernard Teston

► **To cite this version:**

Philippe Abbeg, Robert Espesser, Bernard Teston. Couplage et exploitation d'un synthétiseur à formants sur un ordinateur. Travaux interdisciplinaires du Laboratoire Parole et Langage, 1978, 5, pp.141-157. hal-00241537

HAL Id: hal-00241537

<https://hal.science/hal-00241537>

Submitted on 6 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRAVAUX DE L'INSTITUT DE PHONETIQUE D'AIX

Volume 5, 1978. Pp. 141-157

COUPLAGE ET EXPLOITATION
D'UN SYNTHÉTISEUR À FORMANTS SUR UN CALCULATEUR

PHILIPPE ABBEG
(E.S.P. Marseille)

ROBERT ESPESSER & BERNARD TESTON
(Université de Provence)

COUPLAGE ET EXPLOITATION D'UN SYNTHETISEUR

A FORMANTS SUR UN CALCULATEUR

ABBEG P., ESPESSER R. ET TESTON B.

Nous décrivons dans cet exposé l'étude et la réalisation du couplage d'un synthétiseur à formants sur un ordinateur.

Le système se compose des matériels suivants :

- Un ordinateur SEMS T 1600 de 32 K mots de 16 bits équipé des opérations de multiplication, division et opérateur flottant câblés.
- Une unité de disque à cartouche CDC de 5 millions d'octets.
- Une console de visualisation TEKTRONIX 4.012.
- Une imprimante alphanumérique LOGABAX LX 180.
- Un synthétiseur à formants conçu et réalisé au Laboratoire de la Communication parlée de l'Ecole Nationale Supérieure d'Electronique et de Radioélectricité de GRENOBLE (E.N.S.E.R.G.).

DESCRIPTION DU SYNTHETISEUR

Il s'agit d'un synthétiseur à formant à 20 paramètres dont 13 peuvent être contrôlés automatiquement (Tableau 1). Ces derniers sont commandés au moyen d'un dispositif de lecture d'un document, sur lequel sont tracés les évolutions par rapport au temps, des paramètres de la séquence de parole à synthétiser. Ce système qui utilise une caméra de télévision à grande définition est efficace, et très formateur dans son utilisation. Malheureusement, il n'est pas dénué d'inconvénients tels que : manque de précision dans la définition des paramètres, temps de synthèse limité et mal défini, détérioration et dégradation rapide des documents de référence, difficulté d'emploi. En toute connaissance de cause, le constructeur a prévu deux possibilités différentes de commande des paramètres par un ordinateur. Le premier, au moyen de tensions continues, le second au moyen de signaux à rapport cyclique variable agissant directement sur la commande de paramètres.

Dans le premier cas, il est nécessaire de disposer d'un convertisseur numérique analogique et d'un démultiplexeur analogique, dans le second d'un filtrage numérique. Nous n'avons utilisé aucune de ces deux possibilités prévues par le constructeur et en avons imaginé une troisième. Elle consiste à réaliser le couplage au niveau de la commande des paramètres par le système vidéo d'origine, en générant au moyen d'un interface commandé par le ordinateur les signaux ON et OFF.

Physiquement, on retire du synthétiseur les deux tiroirs de commande vidéo et on les remplace par un tiroir unique recevant les informations issues du ordinateur. Ce système a pour principaux avantages :

- de faire contenir tout l'interface dans le synthétiseur sans avoir à réaliser un câblage extérieur.
- d'utiliser les convertisseurs numérique-analogique de chaque paramètre.
- d'utiliser le filtrage analogique des paramètres.
- d'utiliser les dispositifs de cadrage et de commande manuelle de chaque paramètre.

Ces deux dernières possibilités disparaissent si l'on utilise un des deux couplages prévus par le constructeur.

En revanche, on peut reprocher à notre solution, l'inconvénient d'être moins précise mais uniquement par rapport au couplage au moyen des signaux de découpage.

DESCRIPTION DES CIRCUITS DE COUPLAGE

On peut les diviser en deux parties. D'une part, les circuits de

commande des paramètres, d'autre part leurs circuits de calibration (Figure 1).

Les circuits de couplage sont constitués par une horloge (2 MHz) qui incrémente des compteurs binaires (74 93). Lorsque la valeur du comptage est égale au contenu du mot d'information envoyé par le calculateur, on déclenche simultanément les signaux OFF de tous les paramètres, et l'on incrémente au moyen d'un compteur binaire les signaux ON. Ce compteur a un cycle de 16 positions. Les deux premières (0 et 1) servent dans l'ordre à générer les signaux T et RAZ des convertisseurs de commande des paramètres. Les quatorze autres positions (2 à 15) servent à générer les signaux ON de comptage pour quatorze paramètres au choix de l'utilisateur. Ceci laisse la possibilité de faire évoluer le synthétiseur car l'on dispose ainsi d'une voie de commande d'un paramètre supplémentaire.

Les informations provenant du calculateur sont codées sur 8 bits et envoyées au moyen d'un coupleur universel cadencé par une horloge externe. Le synthétiseur en fait l'acquisition au rythme de cette horloge par l'intermédiaire du signal SERVI qui apparaît chaque fois que les informations sont stables à la sortie du coupleur. Il n'y a pas de réponse du synthétiseur car il existe une grande marge de sécurité entre la cadence maximale de sortie des informations du calculateur et la cadence d'utilisation du synthétiseur. Ainsi, il ne s'est pas avéré nécessaire de disposer d'une détection d'erreur de cadence.

Les valeurs des différents paramètres étant ajustables par le contrôle de leurs extremums, nous avons la possibilité de calibrer les paramètres au moyen d'un dispositif manuel d'entrée des informations numériques. L'affichage des niveaux est réalisé par l'intermédiaire de trois roues codeuses DCB dont on peut faire varier les valeurs entre 0 et 255. Après avoir remis à Zéro le compteur d'adresse des paramètres, on peut actionner l'horloge coup par coup. Un afficheur de 2 digits permet de visualiser les paramètres que l'on étalonne. On peut ainsi ajuster leurs valeurs. L'afficheur permet également le contrôle du bon fonctionnement de la commande numérique. Si à la suite d'une séquence de synthèse, il ne revient pas à zéro, c'est que l'on a eu une erreur de transmission.

TRANSFORMATIONS APORTEES AU SYNTHETISEUR

Elles sont peu importantes et au nombre de deux. La première a consisté à changer les valeurs des fréquences de coupure des filtres analogiques afin de pouvoir augmenter la cadence d'apparition des échantillons. La valeur maximale est de 50 Hertz. Afin de pouvoir cadencer le synthétiseur toutes les 5 ms au maximum (dans de très rares applications) soit à une fréquence de 200 Hertz, nous avons divisé la valeur des capacités des filtres par quatre (0,25 et 0,025 Mf). Toutes les valeurs des fréquences de coupure affichées sur la face avant du synthétiseur sont donc à multiplier par quatre.

La seconde a consisté à coder les convertisseurs numériques-analogiques des paramètres sur 8 bits afin d'avoir la résolution maximale. En fait, ce codage sur 256 niveaux n'est nécessaire que pour la commande de la fréquence fondamentale (F_0). Pour les autres paramètres, elle peut paraître excessive, mais a l'avantage d'être unique.

HORLOGE DE CADENCEMENT

Comme nous l'avons déjà signalé, il s'agit d'une horloge externe que le manipulateur utilise pour cadencer la période d'apparition des échantillons de voix synthétique. Le cycle d'échantillonnage comporte 16 positions, il est nécessaire de "battre la mesure" à une fréquence 16 fois supérieures à celle dite "d'échantillonnage". L'utilisateur peut choisir trois périodes d'échantillonnage : 5, 10 et 20 millisecondes, et peut les multiplier par 10, 100 ou 1000 afin de ralentir autant qu'il est nécessaire la séquence synthétisée (Figure 2).

Nous avons également ajouté au synthétiseur un amplificateur de sortie plus "confortable" que celui d'origine pour contrôler les résultats obtenus, ainsi qu'un étage de sortie sur ligne accordée 50 ohms afin de pouvoir distribuer à distance des stimulus synthétiques pour des tests psychoacoustiques essentiellement, sans avoir à transporter le synthétiseur.

DESCRIPTION DU PROGRAMME DE CONTROLE

Ce programme de contrôle, qui permet d'utiliser le synthétiseur sur le calculateur, est un programme conservatoire, rédigé en langage FORTRAN IV. Il occupe un peu moins de 8 K Mots.

L'utilisateur dispose d'un jeu de commandes qui permettent d'une part, de manipuler les données décrivant la séquence à synthétiser, d'autre part, d'écouter cette séquence.

L'essentiel des commandes représente des fonctions d'édition permettant l'élaboration de l'état. Nous appelons "état" un ensemble de valeurs de paramètres du synthétiseur spécifiées à des instants explicitement mentionnés ; c'est l'état qui décrit totalement la séquence à synthétiser.

Après avoir choisi le ou les paramètres qu'il désire traiter, l'utilisateur spécifie les valeurs de ces paramètres aux instants désirés. L'état ainsi construit, se présente sous la forme d'un tableau de 15 colonnes (le temps dans la première colonne et les 14 paramètres mentionnés dans le tableau n° 1), dont chaque ligne représente les valeurs des paramètres mentionnés à l'instant correspondant. Nous conservons par la suite l'appellation "ligne" pour désigner l'ensemble des valeurs des paramètres mentionnées à un instant donné. Les paramètres peuvent être

mentionnés dans un ordre quelconque qui doit ensuite être respecté lors de la spécification des valeurs affectées à ces paramètres. Il est possible par contre, d'entrer les lignes dans un ordre temporel quelconque. On peut ainsi les corriger par surcharge, les supprimer, les translater (après une telle commande, les lignes des instants t_1 à t_2 deviennent les lignes $t_1 + \Delta t$ et $t_2 + \Delta t$, $\Delta t > 0$ ou < 0). La cohérence de telles opérations sur l'état est vérifiée par la détection de lignes inexistantes, et d'éventuels "écrasements" lors de translations temporelles. Le programme autorise des séquences de 70 lignes au maximum.

La figure 3 représente un exemple de dialogue. Il s'agit de certaines modifications apportées à un état antérieurement défini et sauvegardé dans un fichier.

Il est possible d'obtenir à tout instant, un listage de l'état (Figure 4) ; les valeurs listées sont les valeurs réelles des paramètres codées sur 8 bits. La construction de l'état terminée, celui-ci est ré-écrit dans son fichier. Seul ce fichier est connu de l'opérateur.

A partir de cet état, le module de calcul construit le fichier des échantillons. Nous appelons "échantillon" un bloc de 8 mots (16 octets) groupant les 13 paramètres de commande actuellement utilisés, chacun d'eux étant codé sur 8 bits.

Chaque paramètre est calculé toutes les p seconde, où p est la résolution temporelle (le pas) adoptée, par interpolation linéaire par rapport au temps entre deux instants consécutifs spécifiés, et relatif à ce paramètre.

Par exemple, sur la Figure 4, A_0 est interpolé successivement de $t = 0$ à $t = 80$, puis de 80 à 180 etc... A_B de $t = 0$ à $t = 280$ etc...

Le temps de calcul est de 3 à 4 fois le temps réel pour un pas de 10 millisecondes. Le fichier des échantillons va piloter le synthétiseur par l'intermédiaire du module de synthèse.

Chaque échantillon est renouvelé à la fréquence $1/p$. Le débit effectif est de $16/p$ octets par seconde soit, 1,6 K Hz pour un pas de 10 millisecondes.

Afin d'éviter un encombrement mémoire prohibitif et une limitation de la durée de la synthèse, le fichier des échantillons est transmis au synthétiseur au moyen d'un système de chaînage de deux tampons de 0,5 K Mots, situés en mémoire centrale. Le basculement de ces tampons est géré par le driver du synthétiseur. Ce mode d'échange autorise sur le T 1600 une fréquence maximale de cadencement de 6 K Hz. Il est compatible avec un système d'exploitation multitache (Système R BOS/D de Télémécanique dans ce cas).

Le calcul des échantillons et les synthèses sont déclenchés par l'opérateur au moyen d'un ordre de commande unique (Figure 3). Il exis-

te une extension de cette commande permettant la répétition de la synthèse (nombre de répétition, durée de la pause entre chaque séquence répétées). Le calcul n'est effectué que si l'état a été modifié depuis la dernière demande. Une commande permet de charger le pas temporel, les autres commandes ou éléments de dialogues concernent la gestion des fichiers états.

CONCLUSION

A la différence de nombreuses réalisations équivalentes, notre système n'utilise pas d'entrée/sorties graphiques, ceci pour deux raisons :

- Nous ne disposons pas pour l'instant de dispositif d'entrée graphique, or, la visualisation graphique des paramètres n'est intéressante que si l'on a la possibilité de les modifier par la même voie.
- Le système de synthèse paramétrique est utilisé actuellement pour générer des séquences très précisément contrôlées. De ce point de vue, l'entrée graphique peut poser certains problèmes de résolution et de précision. Nous envisageons cependant ce mode de dialogue à moyen terme.

Dans la même optique de précision du contrôle, les initiatives du module de calcul se limitent à des interpolations linéaires. L'élaboration de transitions plus sophistiquées restent à la charge de l'utilisateur, à la différence, par exemple, du système décrit par FERRERO et Al, où le programme de contrôle impose un type de transition en

$$e^{-\frac{t}{T}}$$

Malgré le handicap lié à l'entrée des paramètres exclusivement numérique à système se révèle satisfaisant et a pu être employé sans difficulté par des utilisateurs peu sensibilisés à l'informatique.

PARAMETRE	SYMBOLE	DYNAMIQUE		UNITE
fondamental	F0	50	500	Hz
formant 1	F1	100	1000	
formant 2	F2	500	3000	
formant 3	F3	1000	4500	
amplitude de voisement	A0	0	40	dB
amplitude de nasalité	AN	0	40	
bande de bruit 1	B1	1000	6000	Hz
bande de bruit 2	B2	2000	9500	
amplitude de bruit	AB	0	40	dB
antiformant de bruit	B0	1000	6000	Hz
amplitude de B0	BV	0	40	dB
largeur de F1	D1	40	100	Hz
largeur de F2	D2	50	150	
extension	XX			

Tableau 1 :
Description des Paramètres

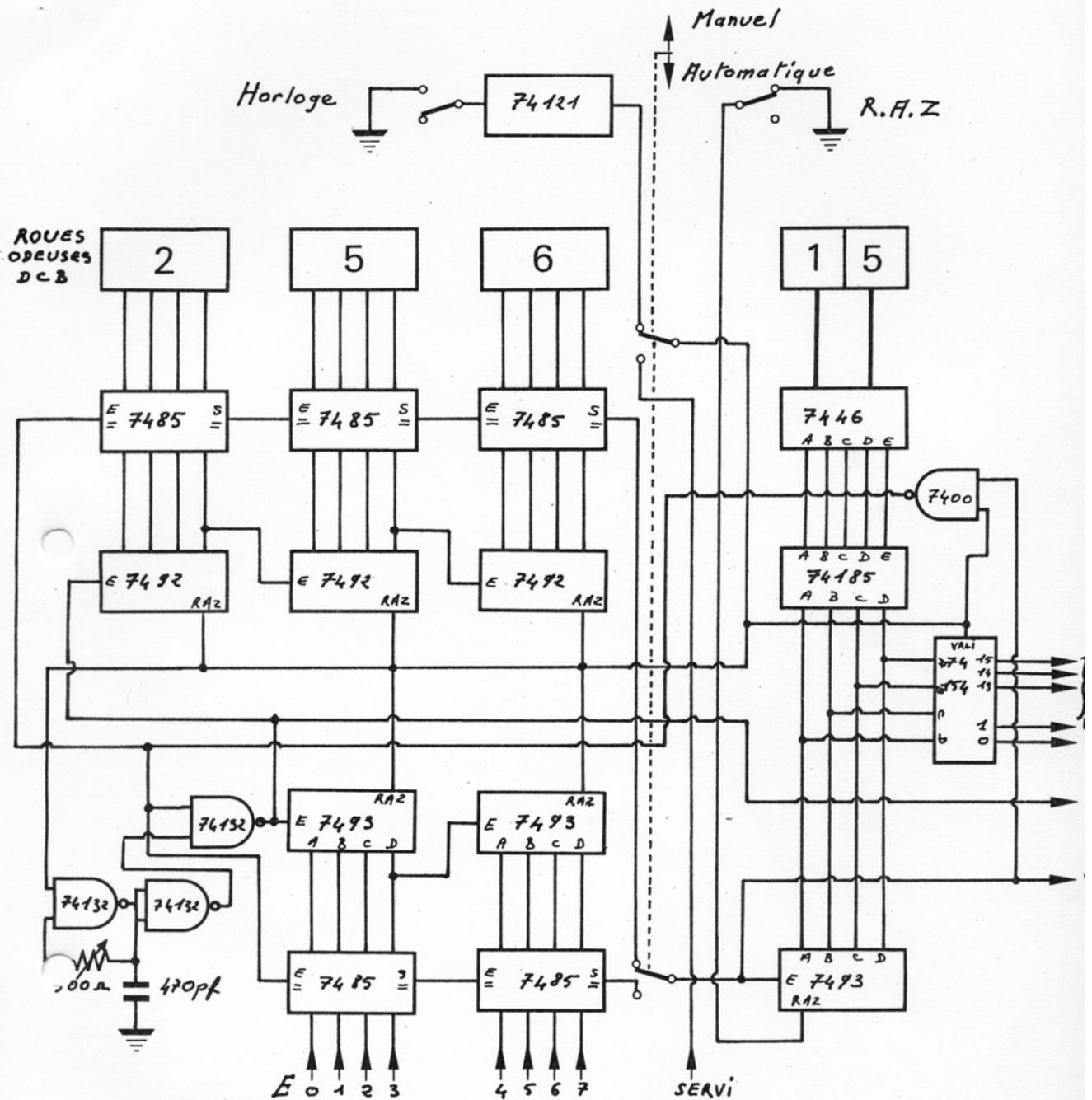


Figure 1

Schéma de l'interface calculateur-synthétiseur

BIBLIOGRAPHIE

- CHAFCOULOFF M., (1976)
Vingt cinq années de recherches en synthèse de la parole.
Editions du CNRS, Paris, p. 287.
- E.N.S.E.R.G., (1974)
Synthétiseur à formants : description et exploitation.
- FERRERO F., VAGGES, RIGHINI G., PELAMATTI M., (1976)
Una sistema di sintesi dell'italiano : primi risultati.
Centro di studio per le ricerche di fonetica (CNR) Padova, p. 25.
- FLANAGAN J.L., (1972)
Speech analysis synthesis and perception.
Springer Verlag, Berlin, p. 444.
- M'RAYATI M., (1971)
Interface Synthétiseur-Ordinateur.
Rapport D.E.A. ENSERG GRENOBLE
- M'RAYATI M., (1973)
Commande d'un synthétiseur à formants par ordinateur.
Bulletin de l'Institut de Phonétique de Grenoble, vol. II, pp. 127-137.