

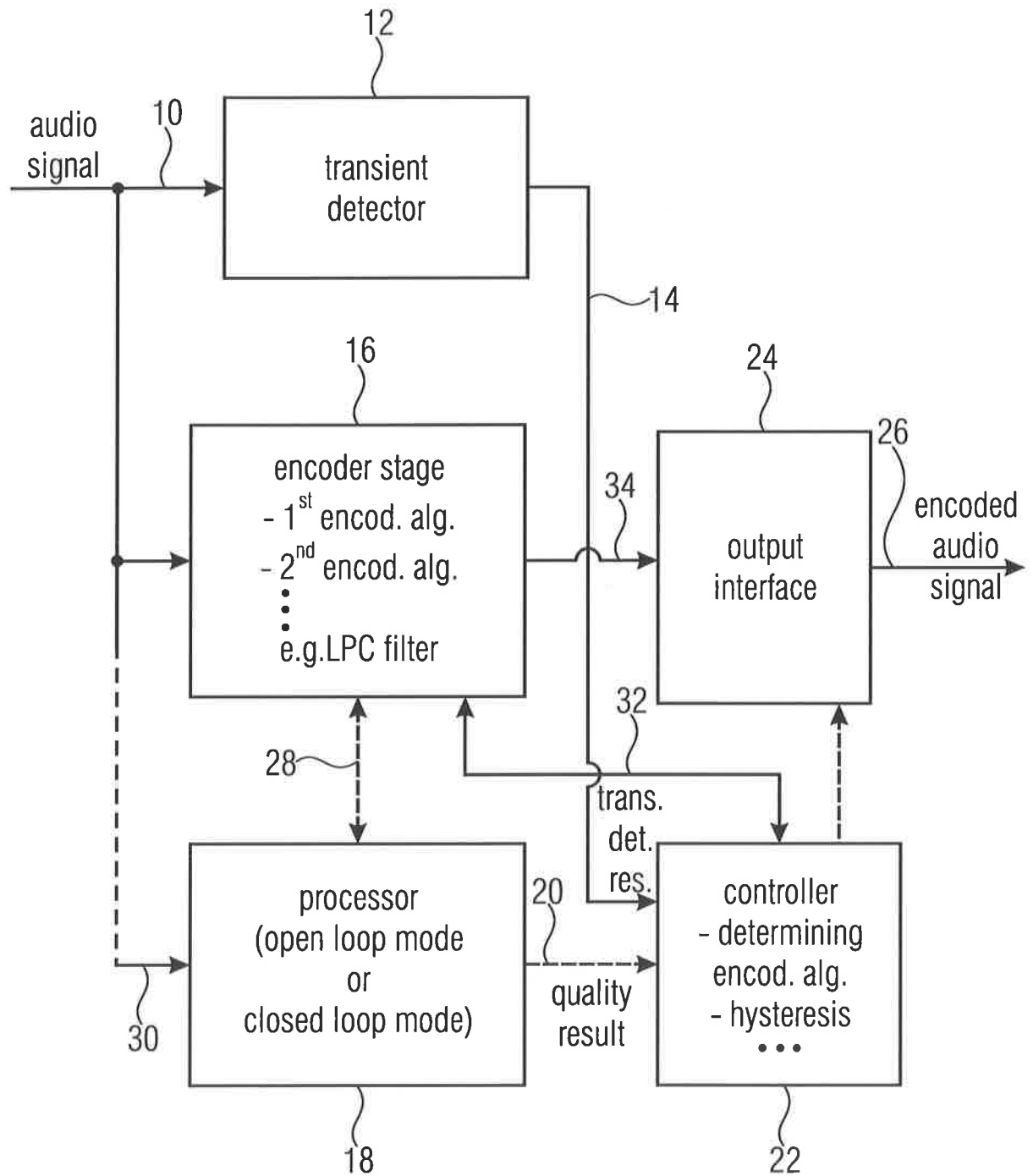


FIGURE 1

encoding algorithm	transient signals	non-transient signals
ACELP	X	
TCX 20		X

FIGURE 2

ACELP ——— TCX20 switch-over

1. Quality Condition

If quality condition indicates
high quality distance (e.g. more than 1dB SNR)
switch over to higher quality encoding algorithm.

2. Transient Condition

If quality condition indicates
small quality distance (e.g. 1 or less dB SNR)
switch over to lower quality encoding algorithm,
when transient detection result indicates that lower quality
encoding algorithm fits to the audio signal characteristic
otherwise switch to higher quality encoding algorithm.

3. Hysteresis condition

Only switch to lower quality encoding algorithm,
when less than the last N frames have been encoded
with the other algorithm (e.g. $N=5$).

FIGURE 3

earlier frames					
> N TCX	0	X	0	0	X
> N ACELP	X	0	X	0	0
< N TCX	0	X	X	X	0
< N ACELP	X	0	0	0	X
	high quality distance for TCX	high quality distance for ACELP	low quality distance for TCX and transient		low quality distance for ACELP and non-transient

X: switch-over
0: no switch-over

FIGURE 4

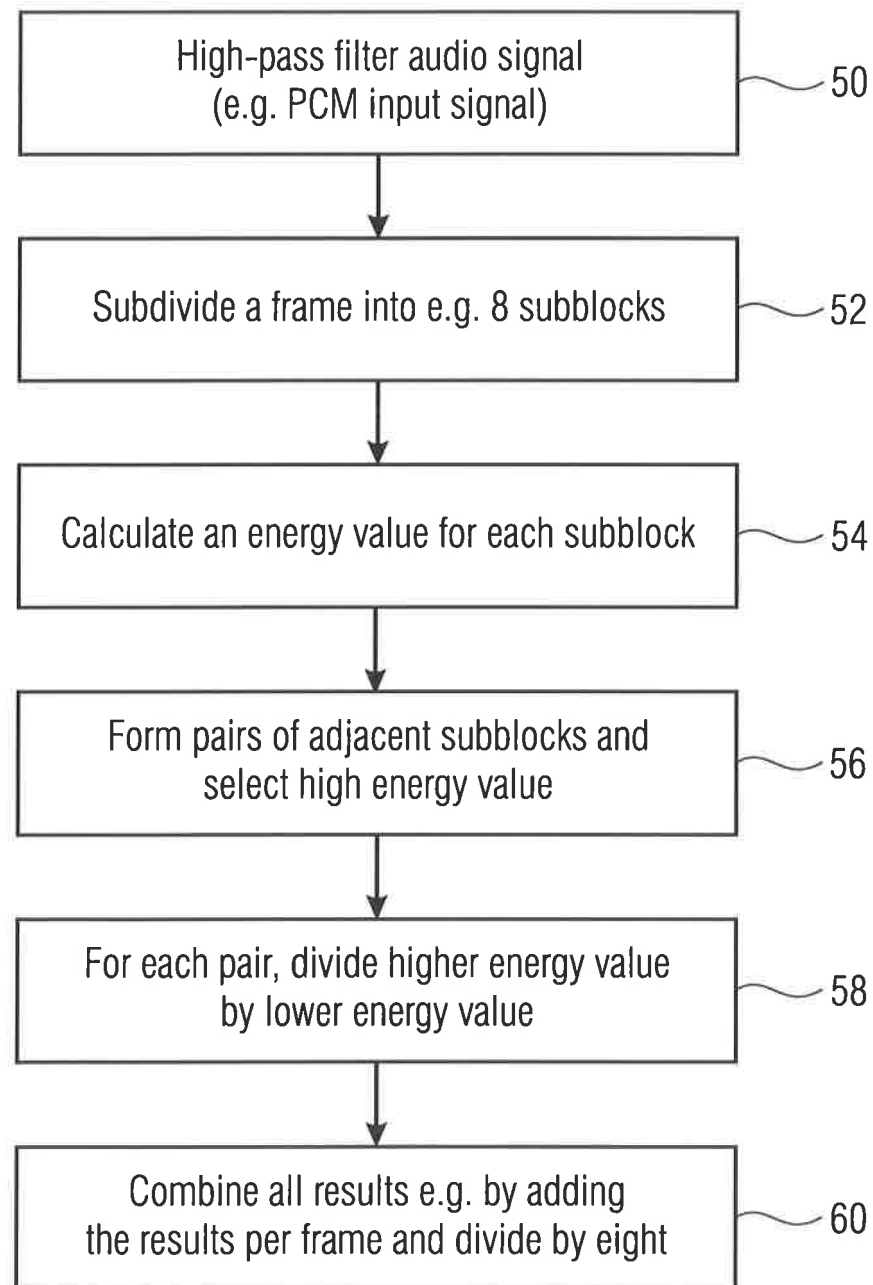


FIGURE 5

6/7

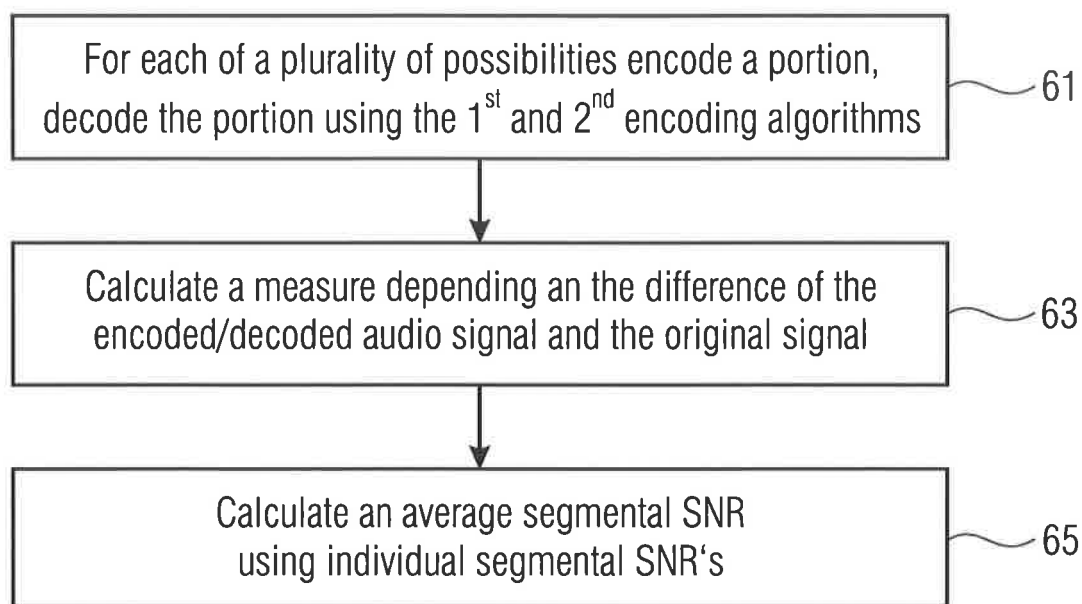


FIGURE 6A

$$\text{segSNR}_i = 20 \log_{10} \left(\frac{\sum_{n=0}^{N-1} x_w^2(n)}{\sum_{n=0}^{N-1} (x_w(n) - \bar{x}_w(n))^2} \right)$$

x_w : wighted audio signal
 $\bar{x}_w$: encoded/decoded weighted signal

$$\overline{\text{segSNR}} = \frac{1}{N_{\text{SF}}} \sum_{i=0}^{N_{\text{SF}}-1} \text{segSNR}_i$$

1 superframe has four frames;
 (1024) (256)

1 frame has four subframes;
 (256) (64)

N_{SF} = number of
 subframes in a frame

i = sample number index

$N = 64 \hat{=}$ number of
 samples in a subframe

FIGURE 6B

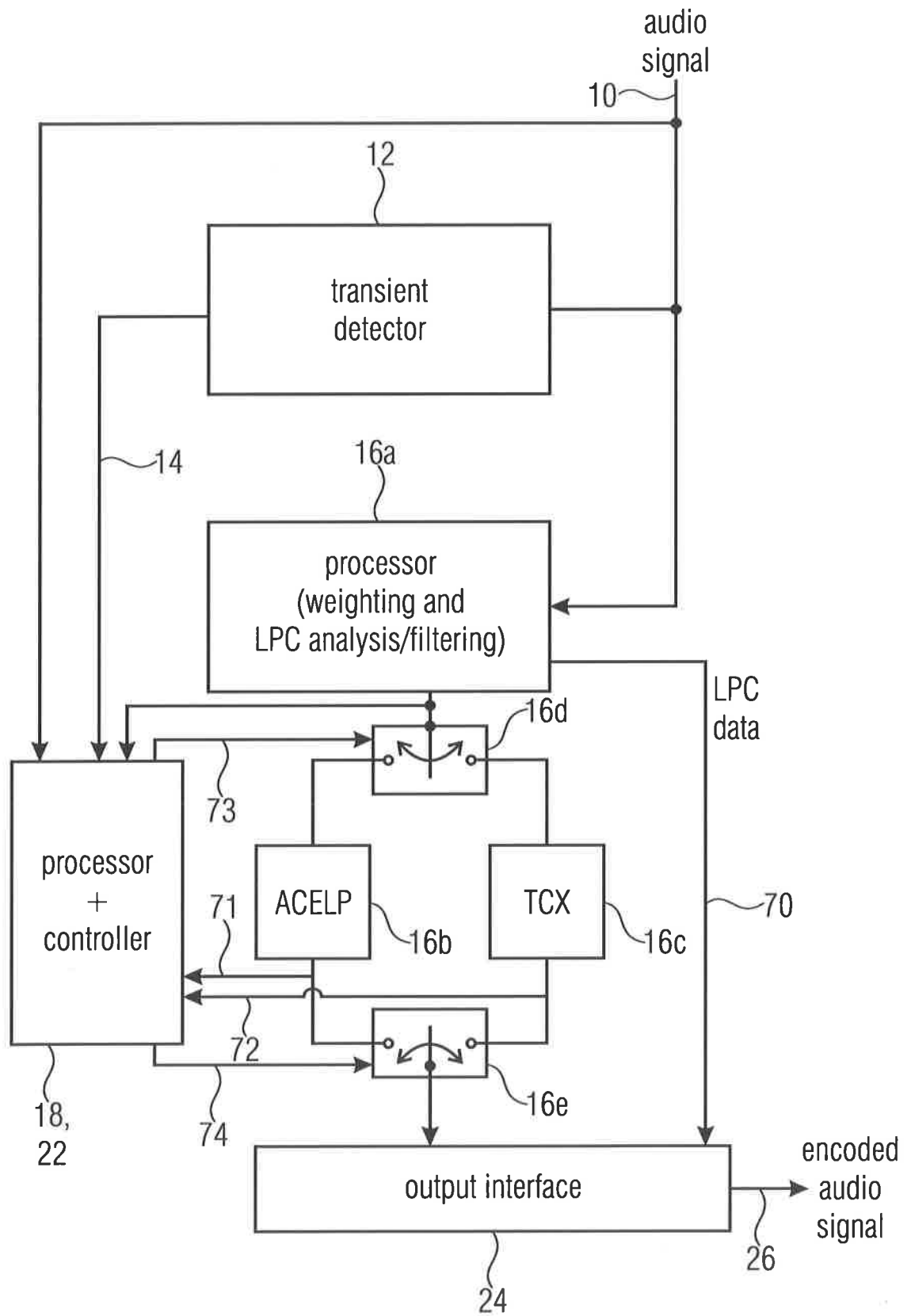


FIGURE 7