

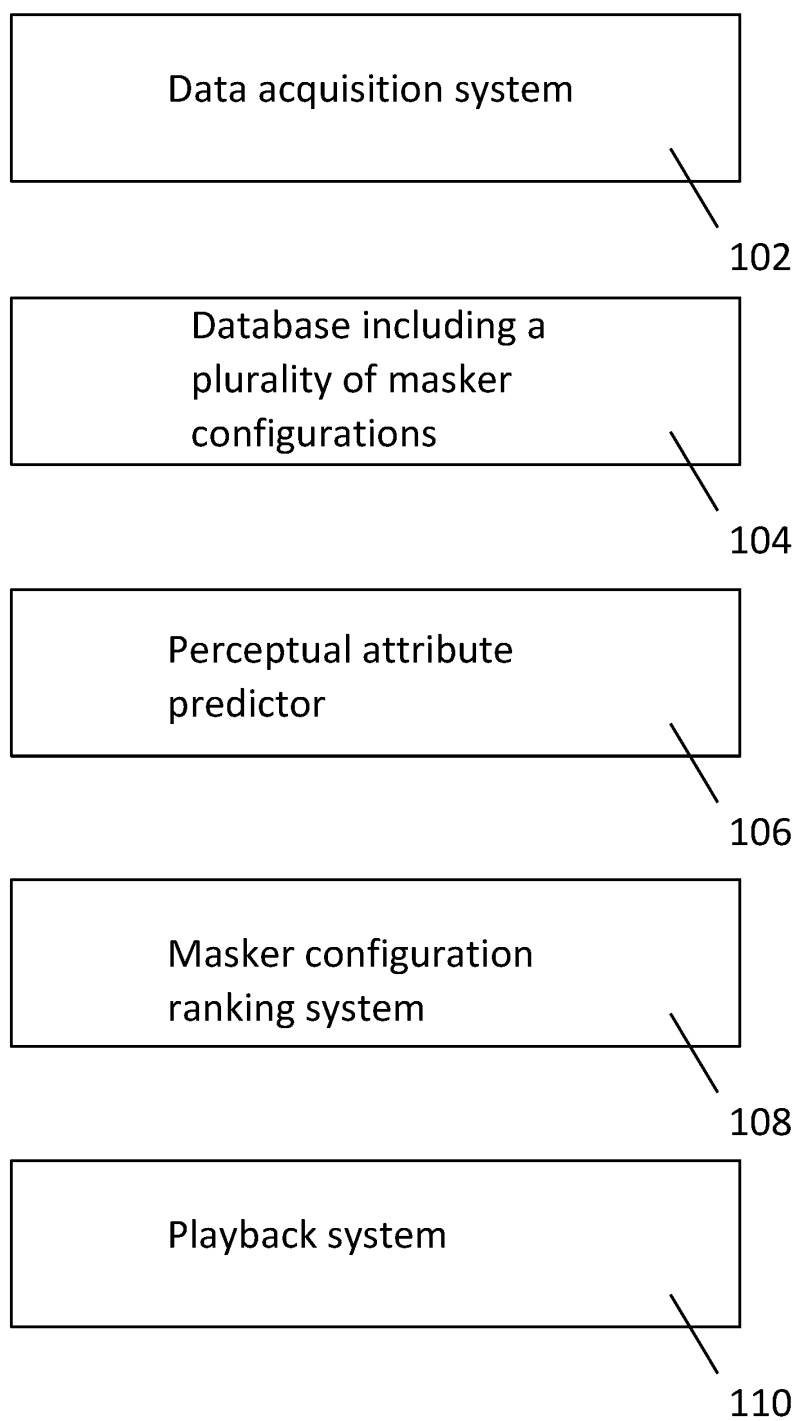
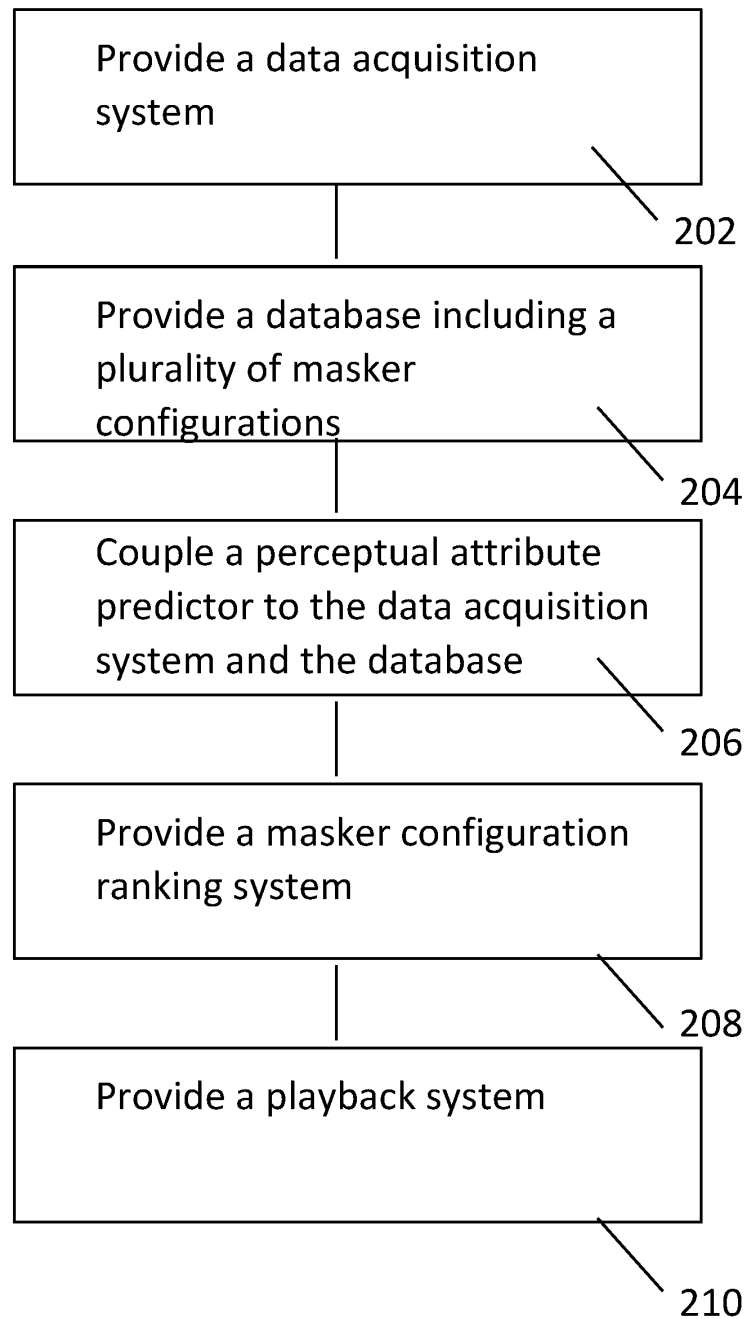
FIG. 1

FIG. 2

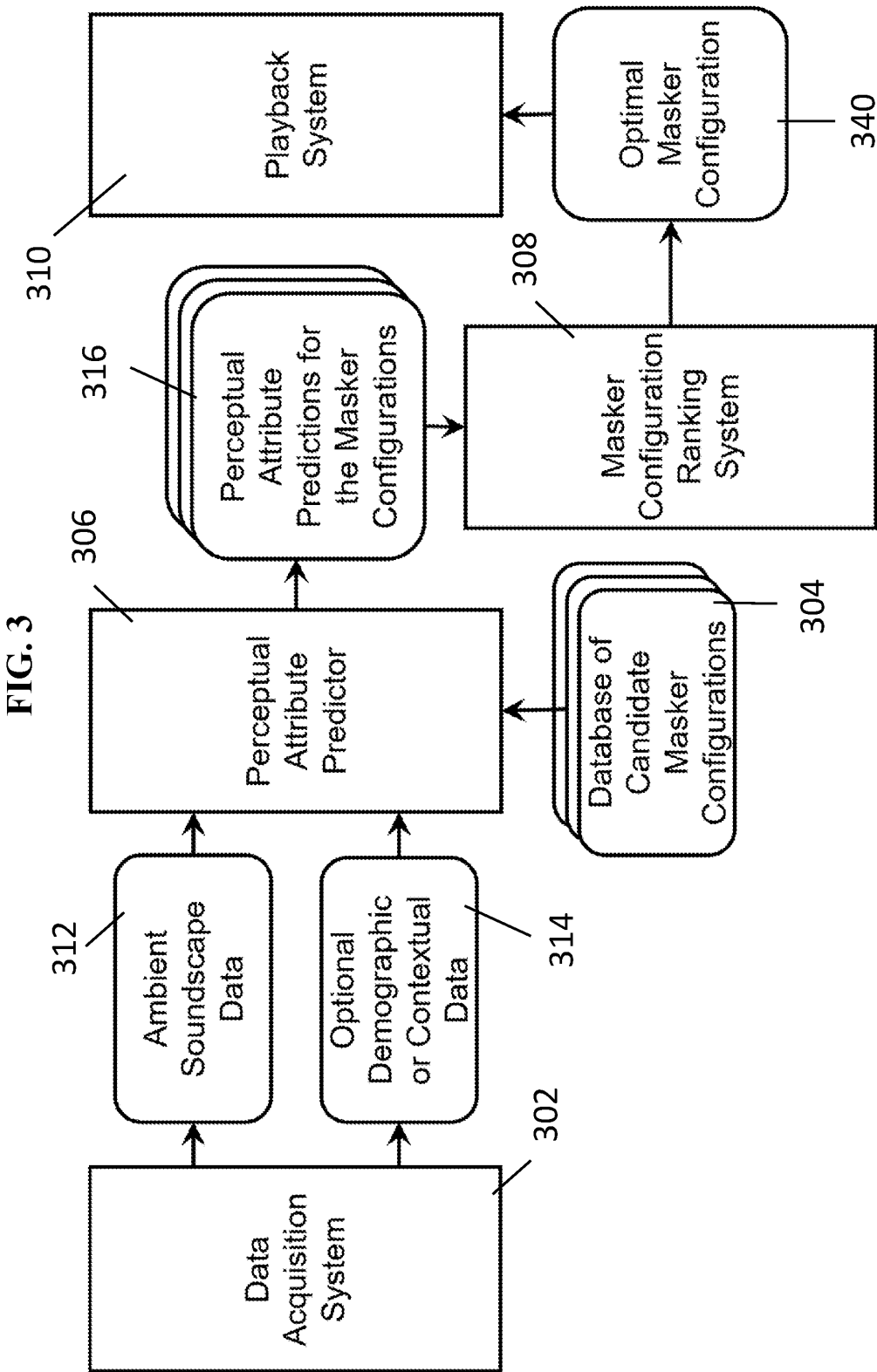


FIG. 4

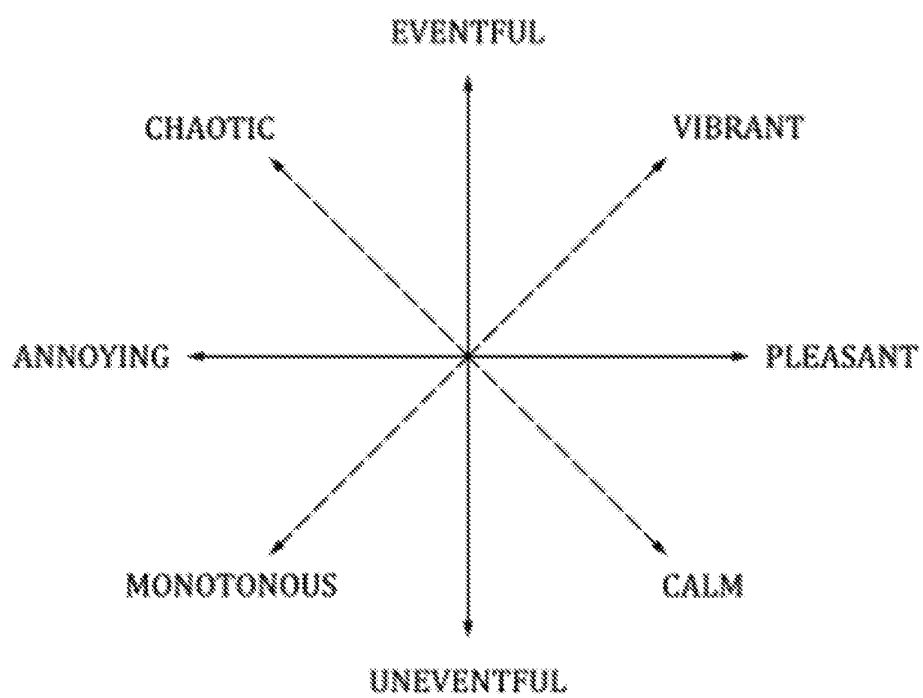


FIG. 5

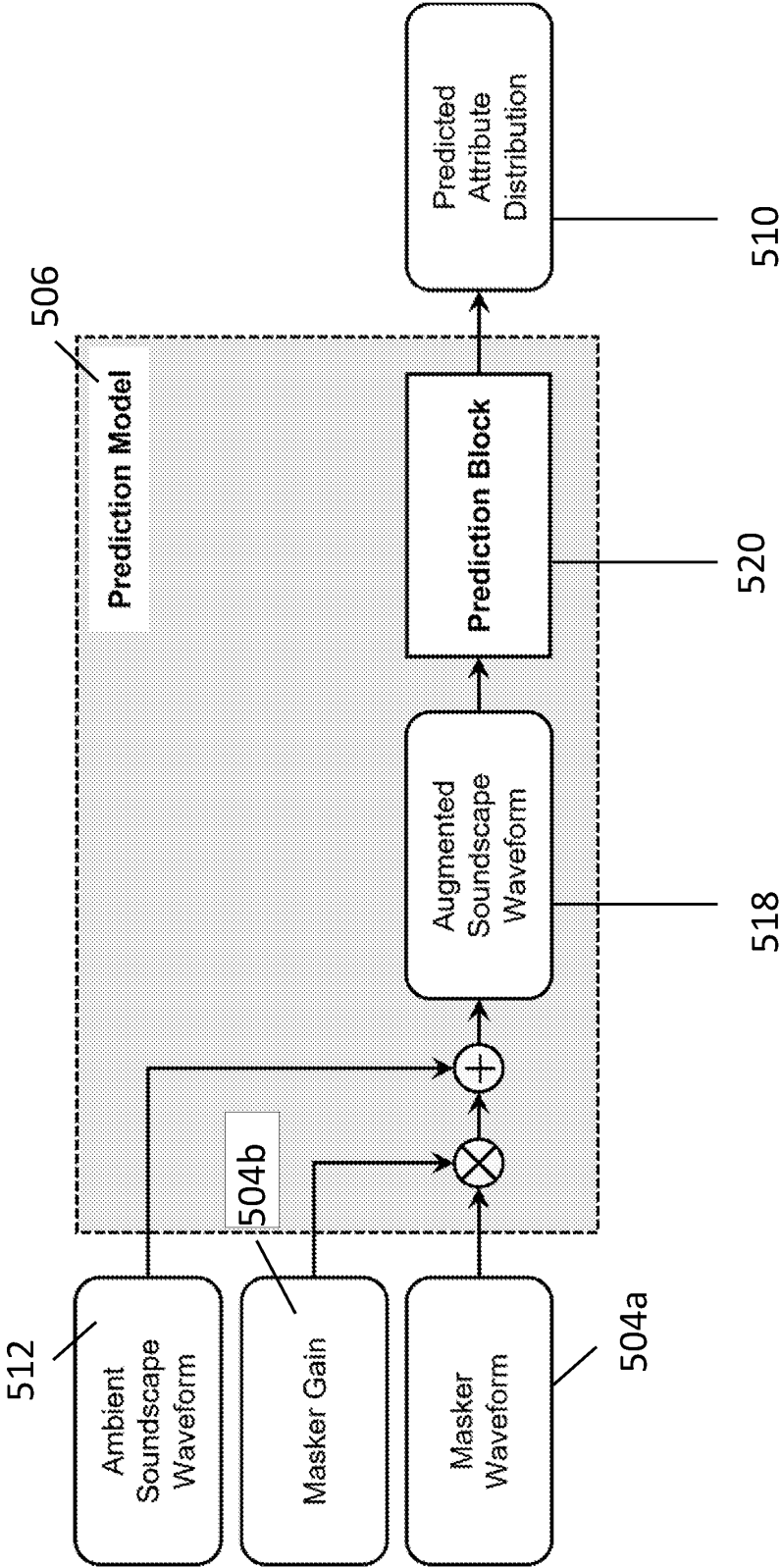


FIG. 6

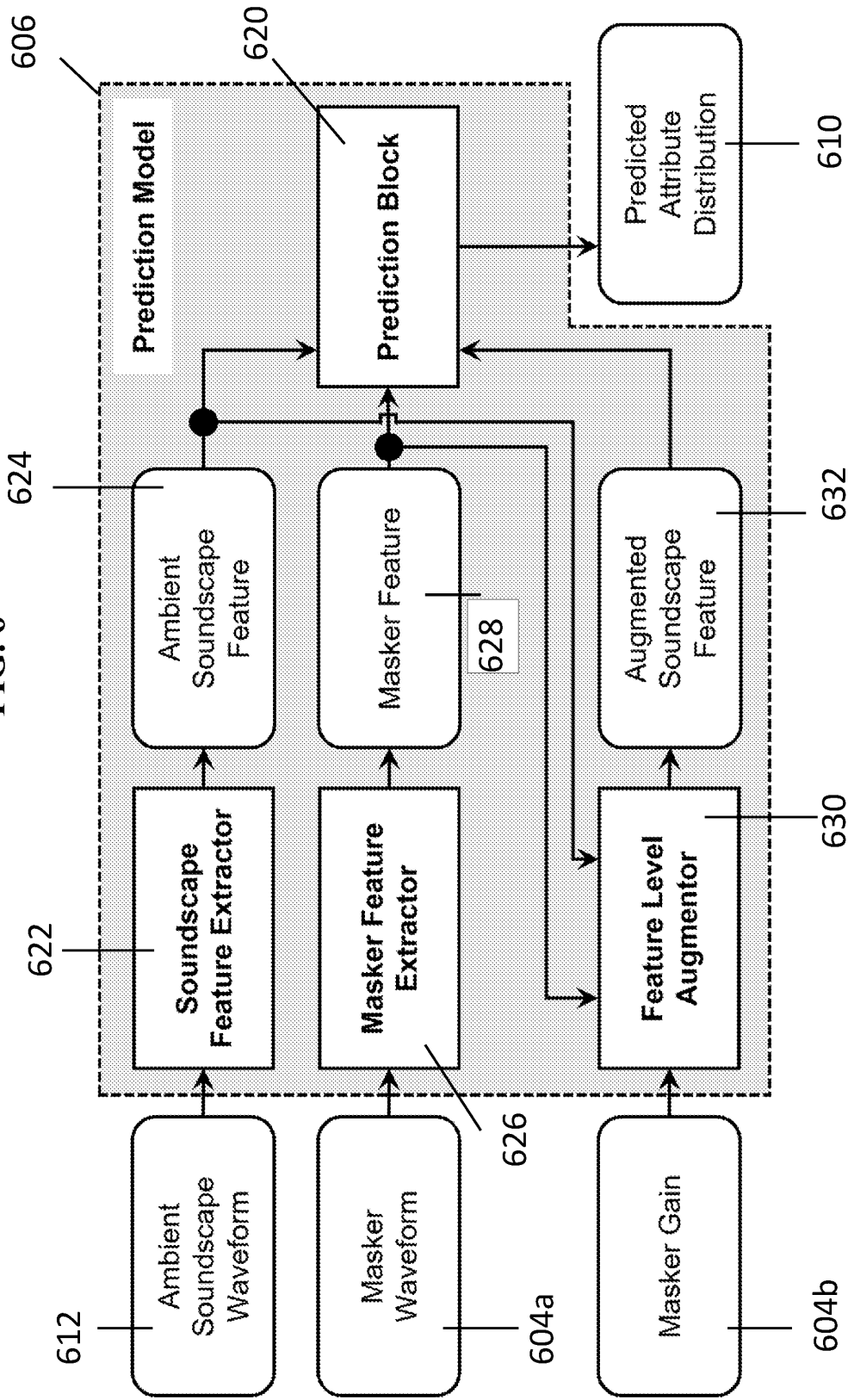


FIG. 7

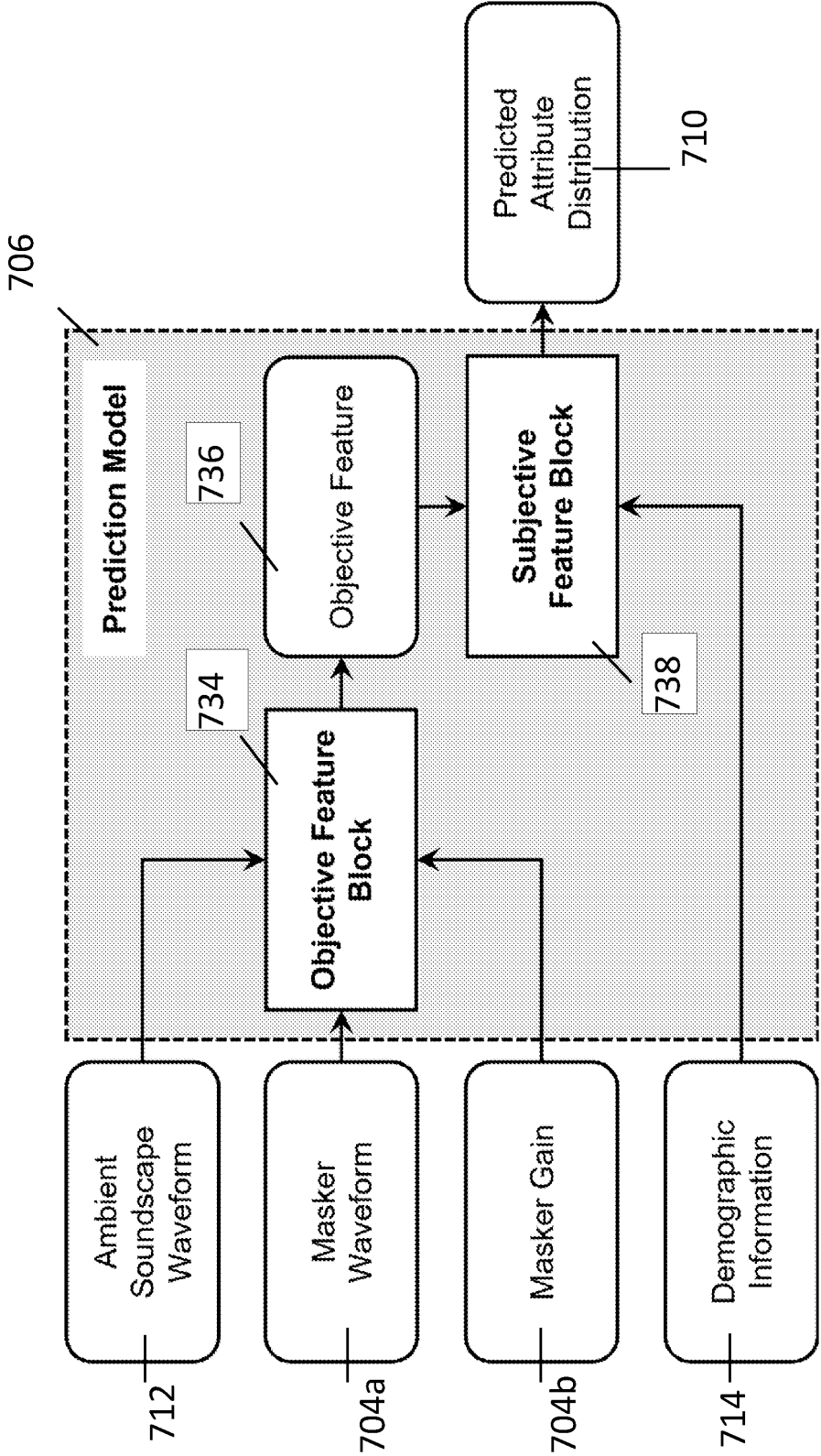


FIG. 8

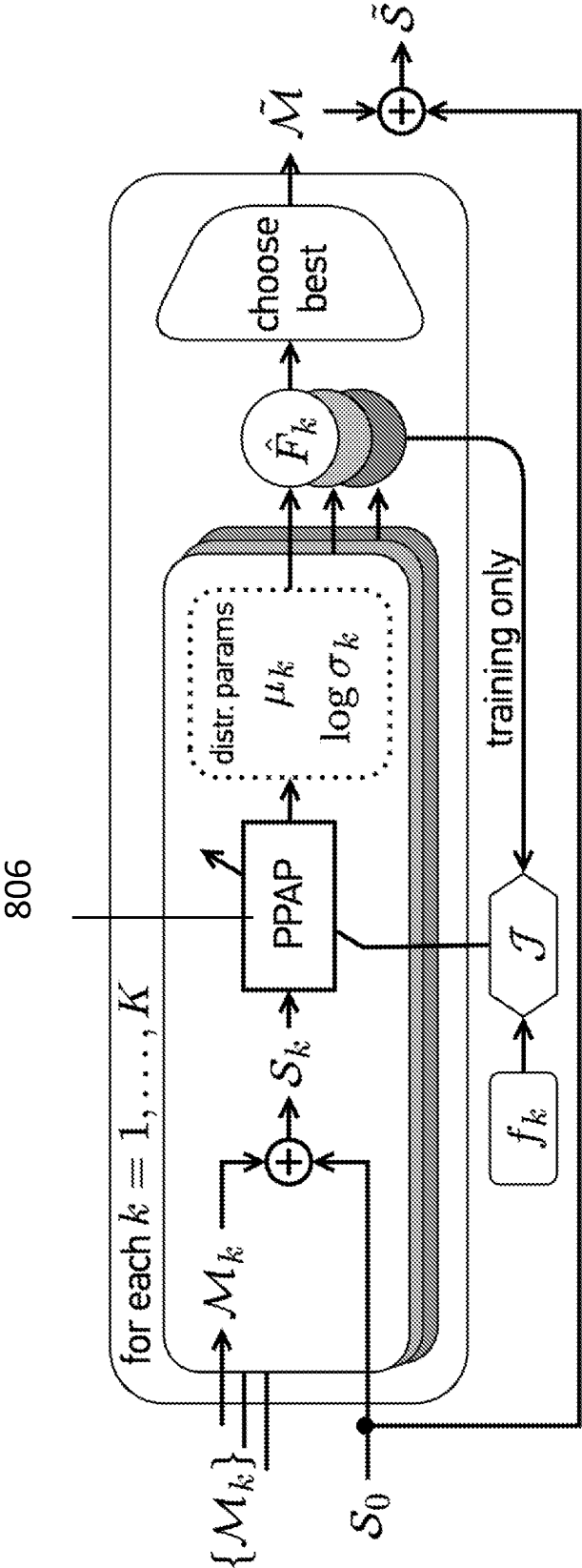


FIG. 9

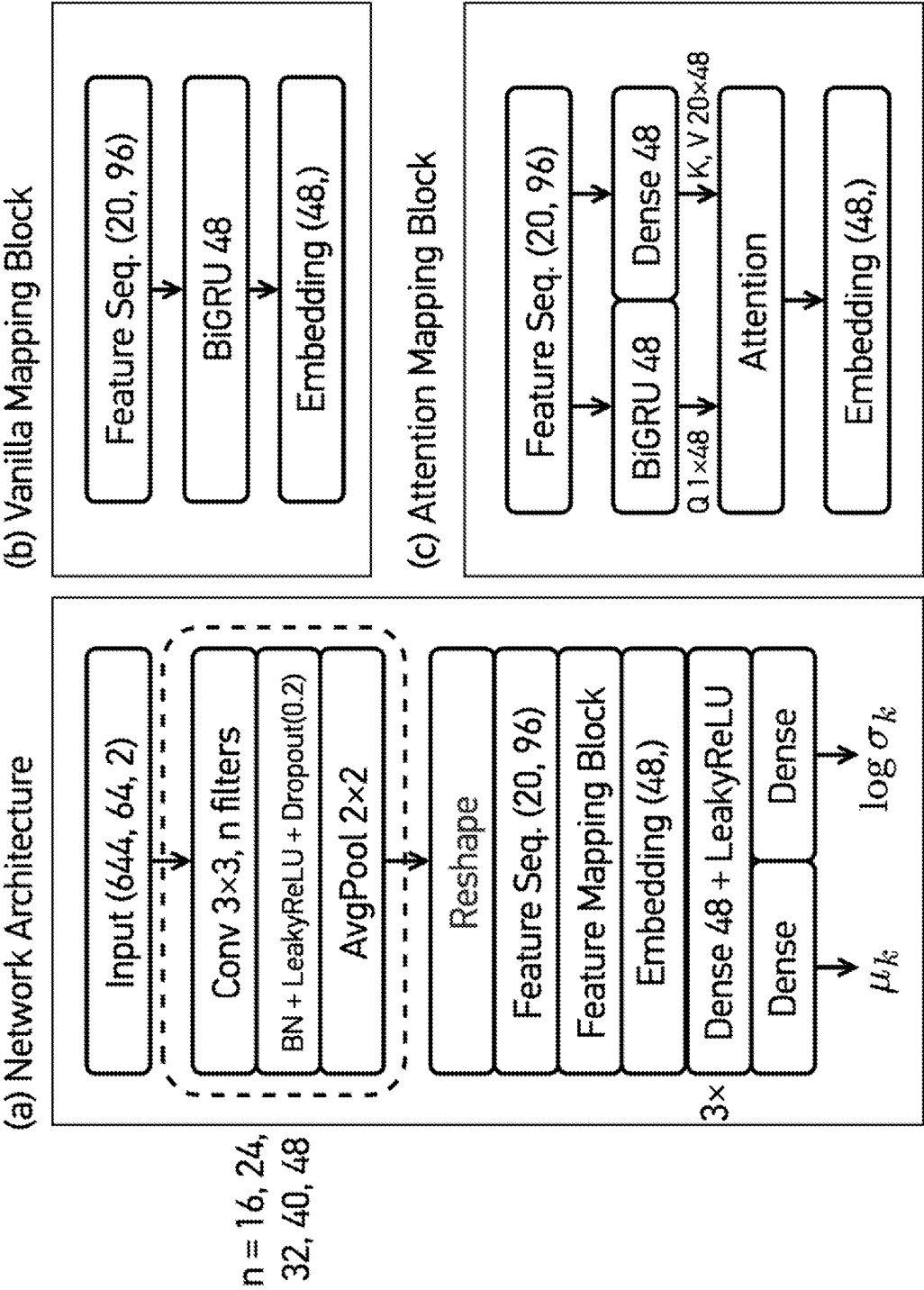


FIG. 10

Cross-Validation Set				
Model	Params.	Deterministic	Probabilistic	Improv. (%)
(Label mean)	—	0.1556	—	—
PPAP w/o attention	76K	0.1262±0.0009	0.1255±0.0011	0.5
PPAP w/ additive att.	85K	0.1264±0.0015	0.1251±0.0014	1.0*
PPAP w/ dot-prod. att.	85K	0.1259±0.0012	0.1250±0.0013	0.7*
PPAP w/ multi-head att.	123K	0.1244±0.0009	0.1230±0.0010	1.1*

Test Set				
Model	Params.	Deterministic	Probabilistic	Improv. (%)
(Label mean)	—	0.1181	—	—
PPAP w/o attention	76K	0.1106±0.0114	0.1020±0.0096	7.8*
PPAP w/ additive att.	85K	0.1030±0.0137	0.0970±0.0080	5.8*
PPAP w/ dot-prod. att.	85K	0.0950±0.0131	0.0906±0.0091	4.6
PPAP w/ multi-head att.	123K	0.0999±0.0147	0.0996±0.0104	0.3

FIG. 11

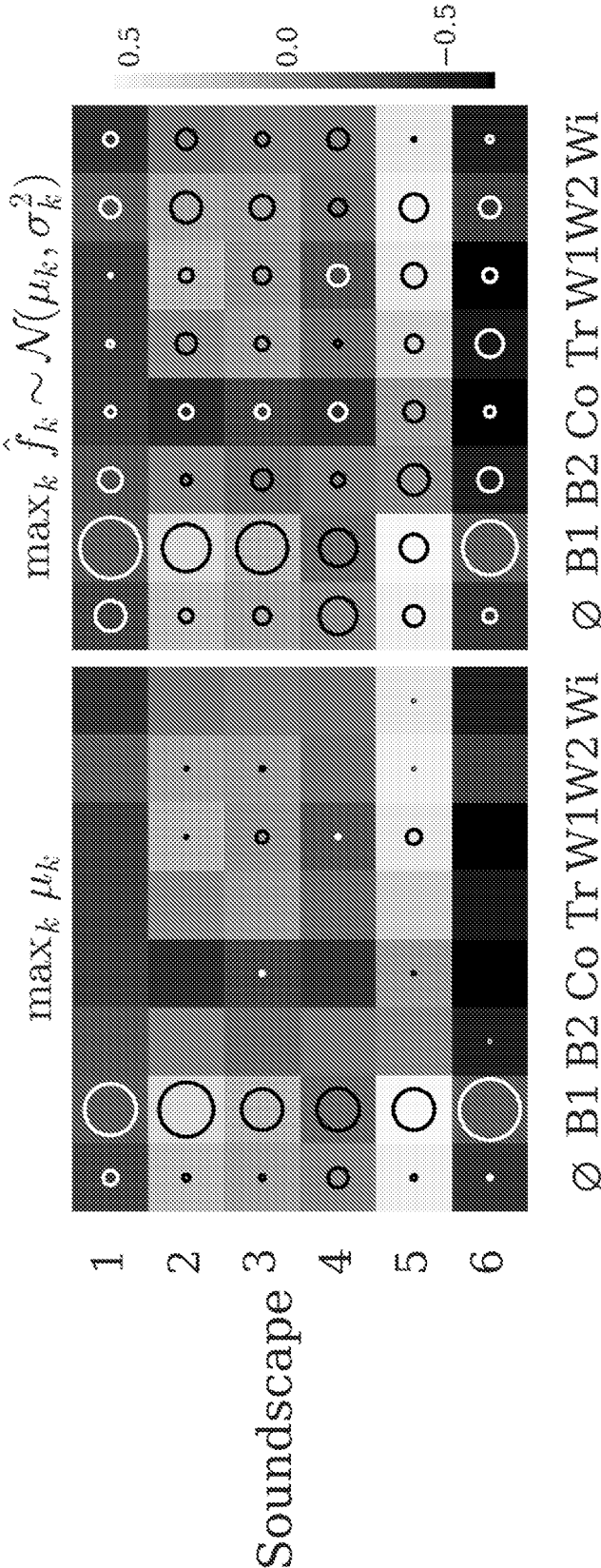


FIG. 12A

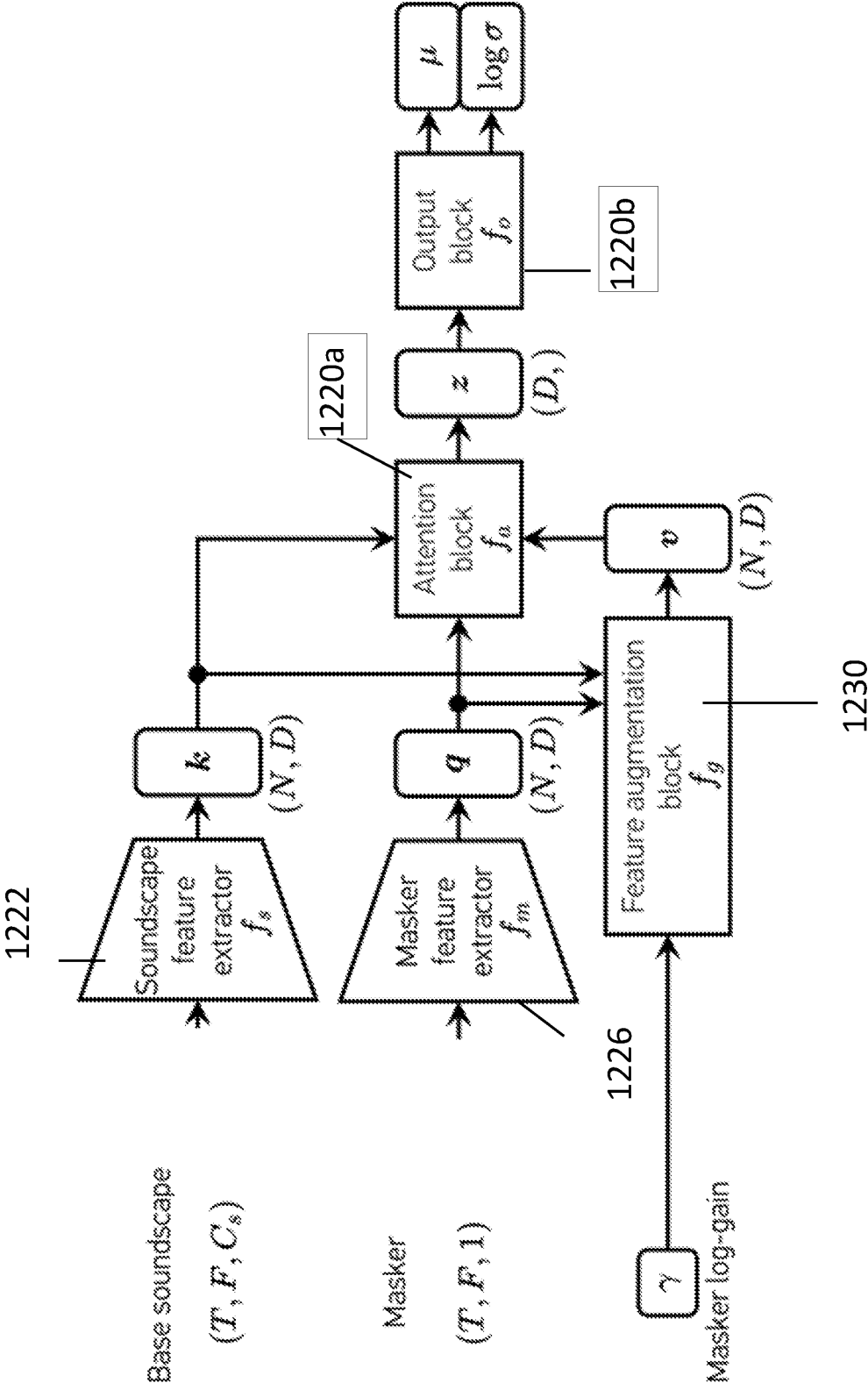


FIG. 12B

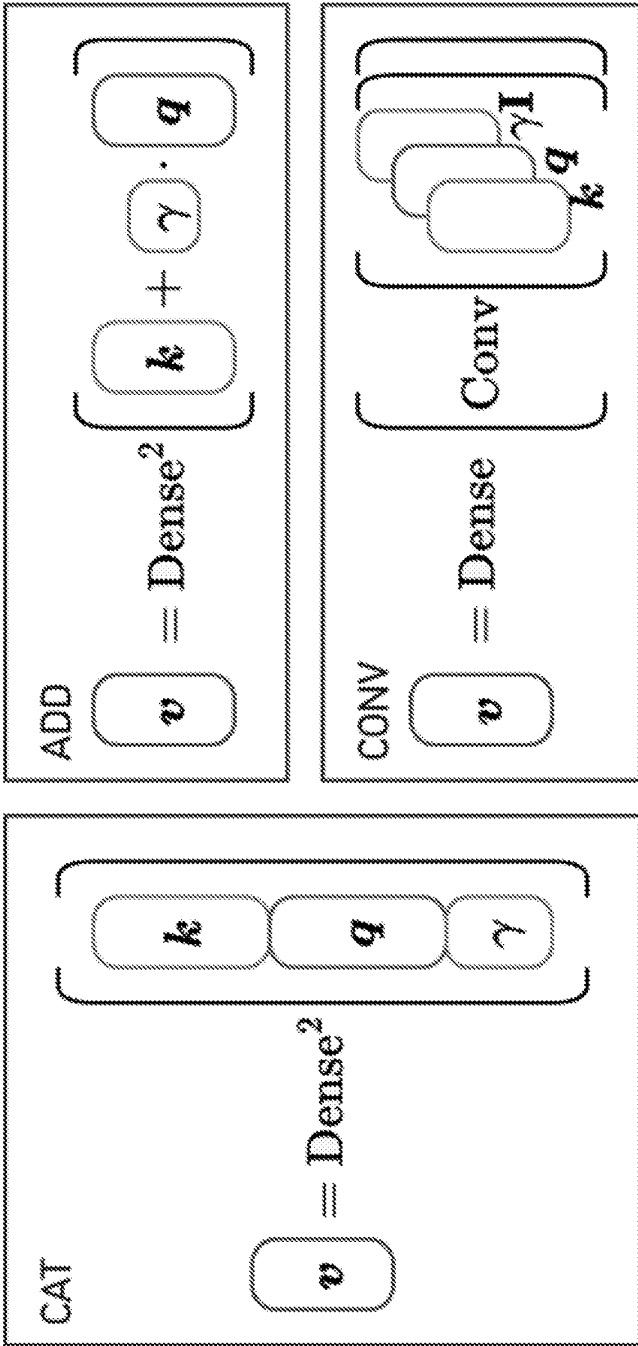


FIG. 13

Algorithm 1: Optimized Inference**Input:** Soundscape recording s_i 1 $\mathbf{S}_i = \text{LogMelSpec}(s_i)$ 2 $\mathbf{k}_i = f_s(\mathbf{S}_i)$ 3 **foreach** masker spectrogram $\mathbf{M}_j$ **do**4 $\mathbf{q}_j = f_m(\mathbf{M}_j)$ // can be precomputed5 **foreach** gain g **do**6 $\mathbf{v}_{i,j,g} = f_g(\mathbf{k}_i, \mathbf{q}_j, \gamma)$ 7 $Y_{i,j,g} = f_o \circ f_a(\mathbf{q}_j, \mathbf{k}_i, \mathbf{v}_{i,j,g})$ 8 **end**9 **end****Output:** Predicted pleasantness distributions $\{Y_{i,j,g}\}$

FIG. 14

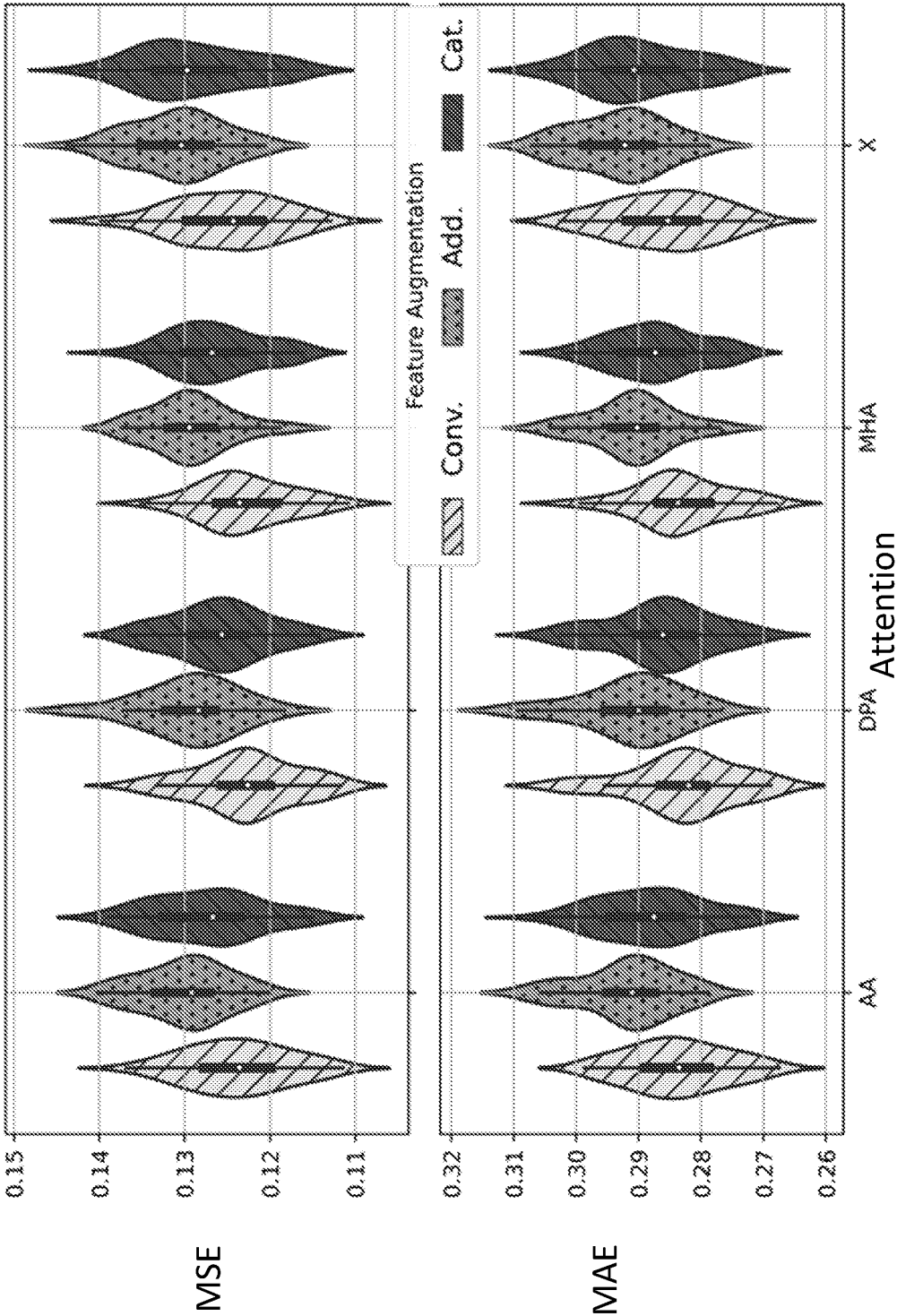


FIG. 15A

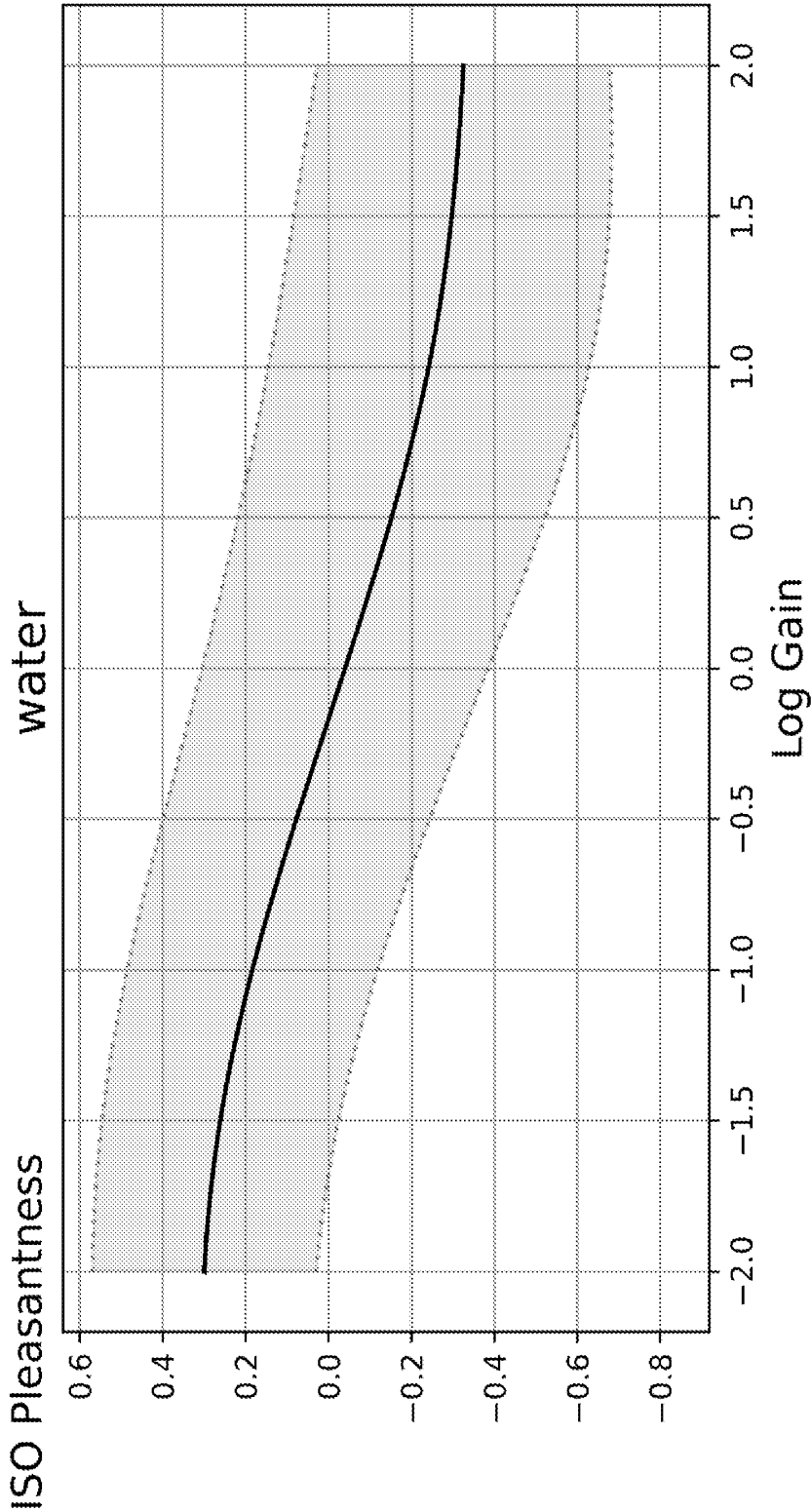


FIG. 15B

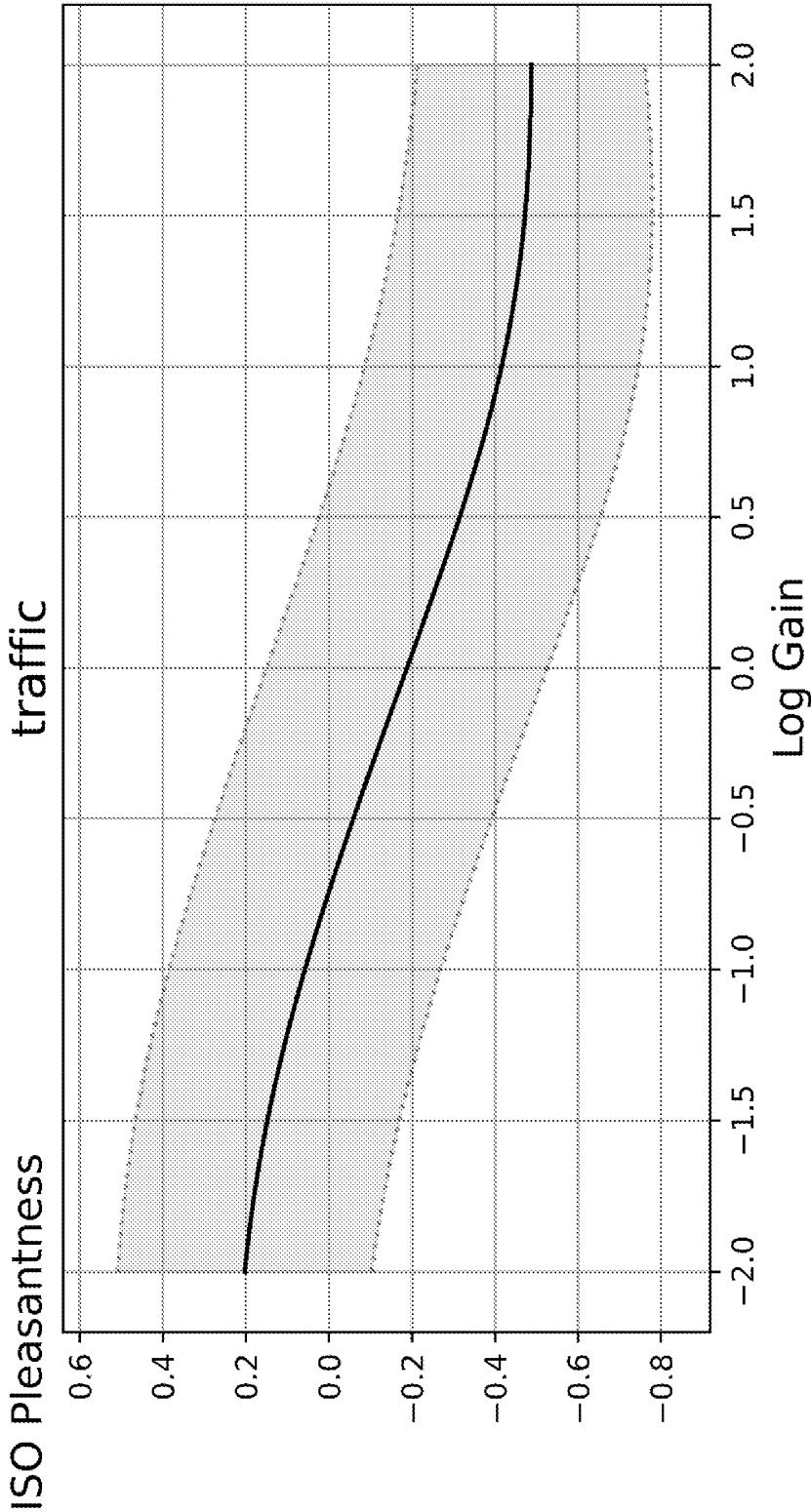


FIG. 15C

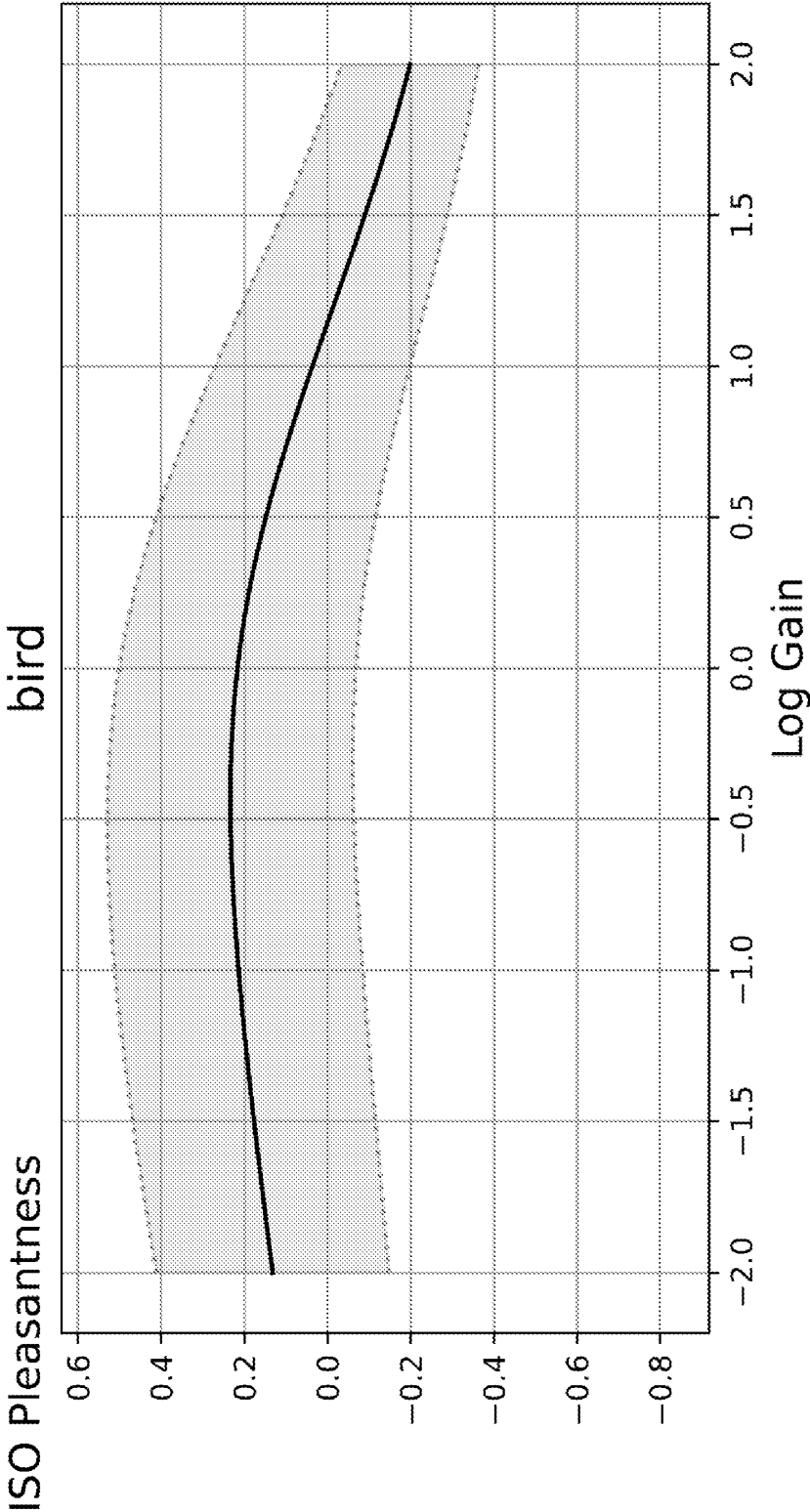


FIG. 15D

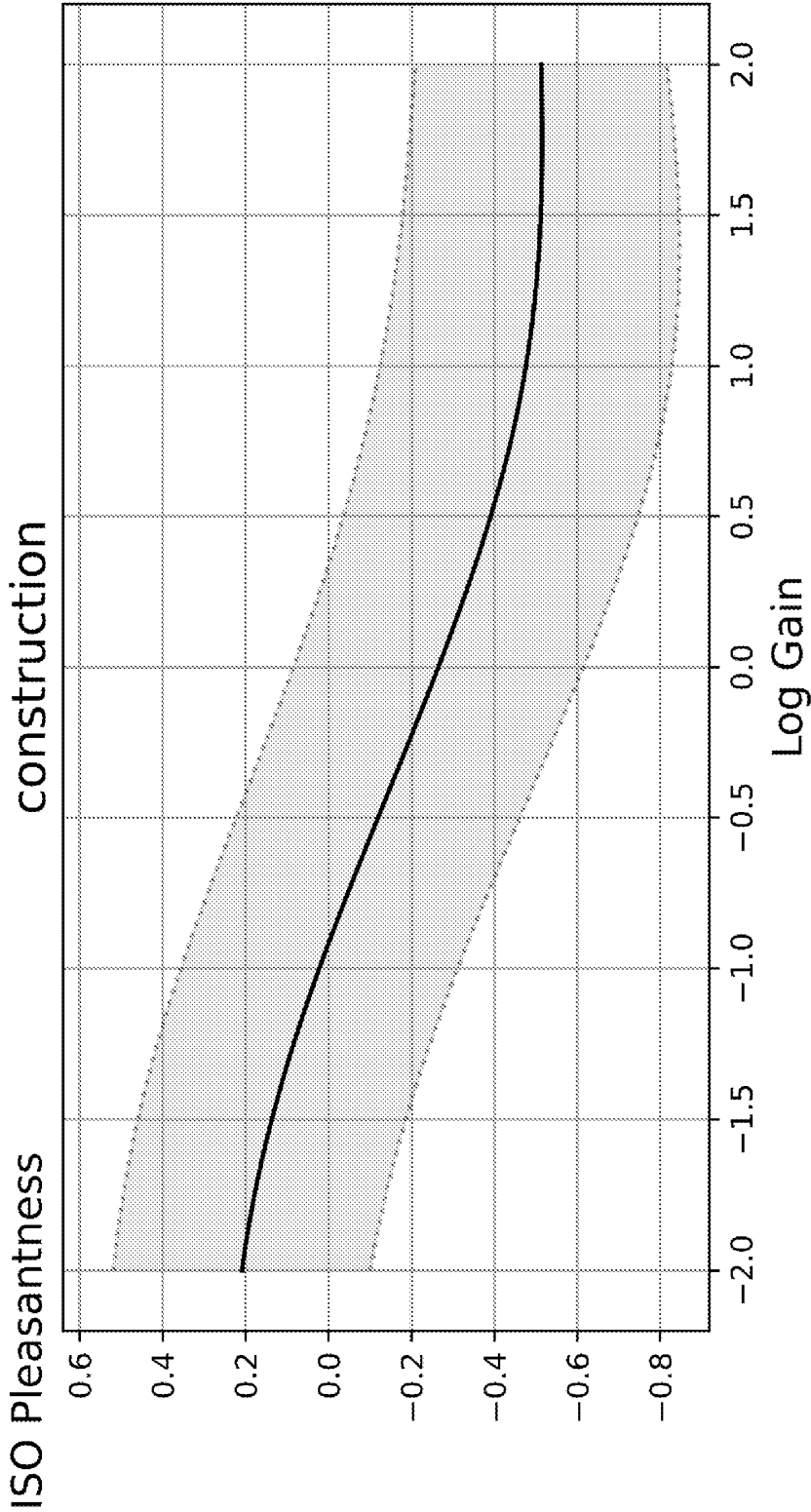


FIG. 15E

