What is claimed is:

1. A method of recognizing a speech command, comprising:

generating a background acoustic model for a sound using a first sound sample, wherein the background acoustic model is characterized by a first precision metric that is less than a first predefined value, and is characterized by a first phoneme precision, and the first sound sample is speech data;

generating a foreground acoustic model for the sound using a second sound sample, wherein the foreground acoustic model is characterized by a second precision metric that is greater than a second predefined value, and is characterized by a second phoneme precision higher than the first phoneme precision;

receiving a third sound sample corresponding to the sound;

decoding the third sound sample by:

using the foreground acoustic model and the background acoustic model, assigning a first weight to the third sound sample corresponding to a probability that the sound sample originated in a foreground;

determining if the first weight meets predefined criteria for assigning the third sound sample to the foreground;

in accordance with a determination that the first weight meets the predefined criteria, interpreting the third sound sample as a portion of a speech command; and

in accordance with a determination that the first weight does not meet the predefined criteria, forgoing recognition of the third sound sample as a portion of a speech command.

2. The method of claim 1, wherein:

the foreground acoustic model for the sound includes a hidden Markov model; and

the background acoustic model for the sound includes a hidden Markov model.

3. The method of any of claims 1-2, wherein:

the first precision metric corresponds to an inverse of a variance of the background acoustic model; and

the second precision metric corresponds to an inverse of a variance of the foreground acoustic model.

4. The method of any of claims 1-3, wherein the foreground acoustic model for the sound and the background acoustic model for the sound comprise a Gaussian mixture model.

5. The method of any of claims 1-4, wherein:

the sound is represented in the background acoustic model as one or more monophones; and

the sound is represented in the foreground acoustic model as one or more triphones.

6. The method of any of claims 1-5, wherein the first predefined value and the second predefined value are equal.

7. The method of any of claims 1-~~6~~5, wherein the first predefined value is less than the second predefined value.

8. The method of any of claims 1-7, further including:

constructing a decoding network having a first path corresponding to the foreground acoustic model and a second path corresponding to the background acoustic model;

assigning a second weight to the third sound sample corresponding to probability that the sound sample originated in a background;

wherein the first weight corresponds to a transfer probability of the third sound sample through the first path and the second weight corresponds to a transfer probability of the third sound sample through the second path.

9. The method of claim 8, wherein the predefined criteria include determining whether the first weight is greater than the second weight.

10. The method of any of claims 1-9, wherein the predefined criteria include determining whether the first weight corresponds to a probability greater than 0.5.

11. An electronic device, comprising:

one or more processors;

a memory; and

one or more programs, wherein the one or more programs are stored in the memory and configured to be executed by the one or more processors, the one or more programs including an operating system and instructions that when executed by the one or more processors cause the electronic device to:

generate a background acoustic model for a sound using a first sound sample, wherein the background acoustic model is characterized by a first precision metric that is less than a first predefined value, and is characterized by a first phoneme precision, and the first sound sample is speech data;

generate a foreground acoustic model for the sound using a second sound sample, wherein the foreground acoustic model is characterized by a second precision metric that is greater than a second predefined value, and is characterized by a second phoneme precision higher than the first phoneme precision;

receive a third sound sample corresponding to the sound;

decode the third sound sample by:

using the foreground acoustic model and the background acoustic model, assigning a first weight to the third sound sample corresponding to a probability that the sound sample originated in a foreground;

determining if the first weight meets predefined criteria for assigning the third sound sample to the foreground;

in accordance with a determination that the first weight meets the predefined criteria, interpreting the third sound sample as a portion of a speech command; and

in accordance with a determination that the first weight does not meet the predefined criteria, forgoing recognition of the third sound sample as a portion of a speech command.

12. The electronic device of claim 11, wherein:

the foreground acoustic model for the sound includes a hidden Markov model; and

the background acoustic model for the sound includes a hidden Markov model.

13. The electronic device of any of claims 11-12, wherein:

the first precision metric corresponds to an inverse of a variance of the background acoustic model; and

the second precision metric corresponds to an inverse of a variance of the foreground acoustic model.

14. The electronic device of any of claims 11-13, wherein the foreground acoustic model for the sound and the background acoustic model for the sound comprise a Gaussian mixture model.

15. The electronic device of any of claims 11-14, wherein:

the sound is represented in the background acoustic model as one or more monophones; and

the sound is represented in the foreground acoustic model as one or more triphones.

16. The electronic device of any of claims 11-15, wherein the first predefined value and the second predefined value are equal.

17. The electronic device of any of claims 11-~~16~~15, wherein the first predefined value is less than the second predefined value.

18. The electronic device of any of claims 11-17, wherein the instructions further cause the device to:

construct a decoding network having a first path corresponding to the foreground acoustic model and a second path corresponding to the background acoustic model;

assign a second weight to the third sound sample corresponding to probability that the sound sample originated in a background;

wherein the first weight corresponds to a transfer probability of the third sound sample through the first path and the second weight corresponds to a transfer probability of the third sound sample through the second path.

19. The electronic device of claim 18, wherein the predefined criteria include determining whether the first weight is greater than the second weight.

20. The electronic device of any of claims 11-19, wherein the predefined criteria include determining whether the first weight corresponds to a probability greater than 0.5.

21. A non-transitory computer readable storage medium storing one or more programs, the one or more programs comprising instructions, which when executed by an electronic device with one or more processors and memory, cause the electronic device to:

generate a background acoustic model for a sound using a first sound sample, wherein the background acoustic model is characterized by a first precision metric that is less than a first predefined value, and is characterized by a first phoneme precision, and the first sound sample is speech data;

generate a foreground acoustic model for the sound using a second sound sample, wherein the foreground acoustic model is characterized by a second precision metric that is greater than a second predefined value, and is characterized by a second phoneme precision higher than the first phoneme precision;

receive a third sound sample corresponding to the sound;

decode the third sound sample by:

using the foreground acoustic model and the background acoustic model, assigning a first weight to the third sound sample corresponding to a probability that the sound sample originated in a foreground;

determining if the first weight meets predefined criteria for assigning the third sound sample to the foreground;

in accordance with a determination that the first weight meets the predefined criteria, interpreting the third sound sample as a portion of a speech command; and

in accordance with a determination that the first weight does not meet the predefined criteria, forgoing recognition of the third sound sample as a portion of a speech command.

22. The non-transitory computer readable storage medium of claim 21, wherein:

the foreground acoustic model for the sound includes a hidden Markov model; and

the background acoustic model for the sound includes a hidden Markov model.

23. The non-transitory computer readable storage medium of any of claims 21-22, wherein:

the first precision metric corresponds to an inverse of a variance of the background acoustic model; and

the second precision metric corresponds to an inverse of a variance of the foreground acoustic model.

24. The non-transitory computer readable storage medium of any of claims 21-23, wherein the foreground acoustic model for the sound and the background acoustic model for the sound comprise a Gaussian mixture model.

25. The non-transitory computer readable storage medium of any of claims 21-24, wherein:

the sound is represented in the background acoustic model as one or more monophones; and

the sound is represented in the foreground acoustic model as one or more triphones.

26. The non-transitory computer readable storage medium of any of claims 21-25, wherein the first predefined value and the second predefined value are equal.

27. The non-transitory computer readable storage medium of any of claims 21-~~26~~25, wherein the first predefined value is less than the second predefined value.

28. The non-transitory computer readable storage medium of claim 21, wherein the instructions further cause the device to:

construct a decoding network having a first path corresponding to the foreground acoustic model and a second path corresponding to the background acoustic model;

assign a second weight to the third sound sample corresponding to probability that the sound sample originated in a background;

wherein the first weight corresponds to a transfer probability of the third sound sample through the first path and the second weight corresponds to a transfer probability of the third sound sample through the second path.

29. The non-transitory computer readable storage medium of claim 28, wherein the predefined criteria include determining whether the first weight is greater than the second weight.

30. The non-transitory computer readable storage medium of any of claims 21-29, wherein the predefined criteria include determining whether the first weight corresponds to a probability greater than 0.5.