We introduce a single channel blind source separation algorithm of audio mixtures. It uses a strategy that is similar to rigid object segregation in videos. A velocity field is defined over the wavelet time-frequency plane. It captures the time evolution of amplitude modulations and harmonic frequencies. Several audio sources are segregated by separating their velocity field with a harmonic rigidity assumption. Signals are then reconstructed from wavelet coefficients in different harmonic templates. The resulting monaural blind source separation is demonstrated on mixtures of speech, singing voice, music, and noise audio signals.