

Ανάλυση εικόνας και μουσικής για την αυτόματη αναγνώριση της διάθεσης



των Δρ. Παναγιώτη Ζέρβα και Δρ. Γιώργου Τριανταφυλλίδη

Ο άνθρωπος λόγω της αλληλεπίδρασης του με τους συνανθρώπους του και το περιβάλλον, καταλαμβάνεται από συναισθήματα και διαθέσεις. Ως διάθεση μπορεί να οριστεί, η εσωτερική υποκειμενική κατάσταση στην οποία μπορεί να βρίσκεται το μυαλό ενός ανθρώπου, και μπορεί να διαρκέσει από ώρες μέχρι μέρες. Έρευνες σχετικά με τη διάθεση και πως αυτή επηρεάζει την συμπεριφορά των ανθρώπων, κατέληξε στο συμπέρασμα ότι είναι μια σχετικά μεγάλης διάρκειας συναισθηματική κατάσταση. Σε αντίθεση με το συναίσθημα, η διάρκεια της διάθεσης είναι μεγαλύτερη με αποτέλεσμα μεμονωμένα γεγονότα ή ερεθίσματα να μην έχουν άμεσο αντίκτυπο στην αλλαγή της. Ενώ συναισθήματα, όπως φόβος ή έκπληξη, μπορούν να αλλιάξουν σε μικρό χρονικό διάστημα, οι διαθέσεις τείνουν να διαρκούν περισσότερο.

Οι διαθέσεις, σε αντιπαράθεση με τα συναισθήματα, δεν σχετίζονται με ένα συγκεκριμένο αντικείμενο. Για παράδειγμα, αν κάποιος διακατέχεται από το συναίσθημα της θλίψης, τότε λέμε ότι είναι λυπημένος για κάποιο συγκεκριμένο γεγονός, όπως μια προσωπική απώλεια. Αντιθέτως, αν κάποιος βρίσκεται σε κακή (λυπημένη) διάθεση, μπορεί να μην είναι για κάτι συγκεκριμένο. Με άλλα λόγια, τα συναισθήματα περιγράφουν έντονες συγκινήσεις που απευθύνονται σε κάποιον ή κάτι, από την άλλη μεριά οι διαθέσεις είναι συναισθηματικές καταστάσεις που τείνουν να είναι λιγότερο έντονες και η εμφάνισή τους δεν είναι προϊόν κάποιων συγκεκριμένων αιτιών.

Για να ξεχωρίσουμε σε αυτό το άρθρο, τη χρήση της έννοιας της διάθεσης με αυτή των συναισθημάτων, υποθέτουμε ότι κάποιο συναίσθημα ουσιαστικά αποτελεί κομμάτι μιας διάθεσης. Διάφορα συναισθήματα μπορούν να εμφανιστούν σε μια διάθεση η οποία αποτελεί την γενικευμένη συναισθηματική κατάσταση του ατόμου.

Σε αυτό το άρθρο, θα μας απασχολήσει ο αυτόματος χαρακτηρισμός διάθεσης σε μια ταινία (πχ μουσικό video clip) με βάση τον ήχο (μουσική επένδυση) και την εικόνα. Και στις δύο περιπτώσεις (μουσική και εικόνα), έχουμε την υπέρθεση συγκινησιακών καταστάσεων με την μορφή διάθεσης, με πολλά συναισθήματα τα οποία μπορεί να αλλιάξουν κατά την χρονική εξέλιξη των γεγονότων.

Αυτόματη Αναγνώριση Διάθεσης σε Ακουστικά Μουσικά Σήματα

Μελέτες που έχουν γίνει στον τομέα της ψυχολογίας, έχουν καταλήξει στην δημιουργία διαφόρων μοντέλων συναισθηματικών καταστάσεων και διαθέσεων, πολλά από τα οποία έχουν χρησιμοποιηθεί με επιτυχία στην κατηγοριοποίηση μουσικών κομματιών. Τα βασικά συναισθήματα τα οποία περιγράφουν σε μεγάλο βαθμό την συναισθηματική κατάσταση ενός ανθρώπου σύμφωνα με τη μελέτη του Ekman είναι θυμός, απδία, φόβος, ευτυχία, λύπη και έκπληξη. Ωστόσο, δεδομένου ότι είχαν μελετηθεί για την κωδικοποίηση των εκφράσεων του προσώπου, κάποιες από τις συναισθηματικές καταστάσεις μπορεί να μην είναι κατάλλη-

δες για την περιγραφή του περιεχομένου μουσικών ακουσμάτων (π.χ. ανδία), επίσης ορισμένες κοινές διαθέσεις που μπορούν να προκύψουν από το άκουσμα μουσικής λείπουν, για παράδειγμα ηρεμία, χαλάρωση κ.α.

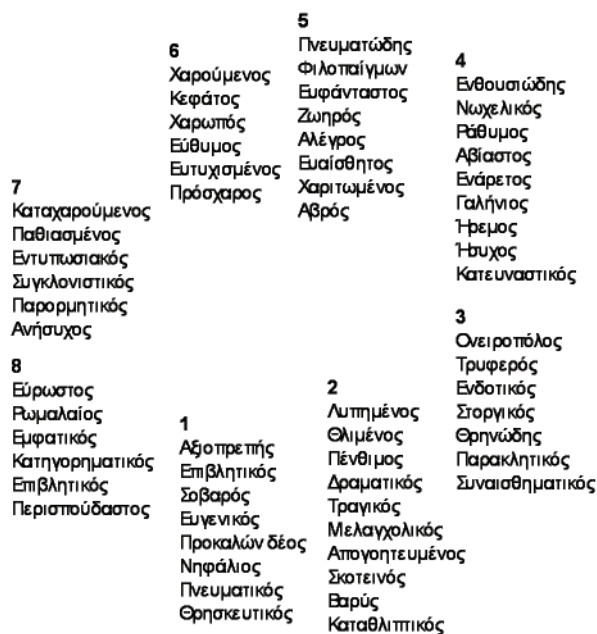
Στην μουσική ψυχολογία, η πιο πρόσφατη και ευρέως διαδεδομένη προσπάθεια δημιουργίας ταξινόμησης σχετικά με την πληροφορία της διάθεσης έγινε από τον Hevner. Ο Hevner σχεδίασε έναν κύκλο ο οποίος αποτελείται από οκτώ ομάδες συναισθημάτων όπως φαίνεται και στην εικόνα 2, από όπου μπορούμε να διακρίνουμε ότι:

- α) τα συναισθήματα, μέσα σε κάθε μια από τις ομάδες έχουν κοντινές σημασίες,
- β) το νοηματικό περιεχόμενο των γειτονικών ομάδων διαφέρει ελάχιστα, και
- γ) η διαφορά μεταξύ ομάδων μεγαλώνει βήμα βήμα, μέχρι να φτάσουμε στην αντιδιαμετρική ομάδα

Τόσο η προσέγγιση του Ekman, όσο και του Hevner, ανήκουν σε αυτά των κατηγορηματικών μοντέλων γιατί το διάστημα των διαθέσεων αποτελείται από ένα σύνολο διακριτών κατηγοριών διάθεσης. Ένα άλλο πολύ γνωστό μοντέλο ταξινόμησης της συγκινησιακής κατάστασης ενός μουσικού κομματιού είναι το IMIRSEL (The International Music Information Retrieval Systems Evaluation Laboratory Project) όπου τα συναισθήματα τοποθετούνται σε ένα συνεχές διαδιάστατο χώρο. Πιο συγκεκριμένα, επιλέγεται ένα σύνολο από 5 συστάδες καταστάσεων διάθεσης, μειώνοντας αποτελεσματικά το «πεδίο τιμών» διάθεσης σε ένα από σύνολο κατηγοριών, που έχουν προκύψει από την μελέτη και επισημείωση μουσικών ακουσμάτων. Για αυτό το λόγο προτείνεται

η χρήση 5 συστάδων διάθεσης, για το πρόβλημα της δημιουργίας μοντέλων μηχανικής μάθησης (machine learning) για την ταξινόμηση μουσικής με βάση το συναισθηματικό φορτίο τους. Οι κατηγορίες αυτές είναι οι παρακάτω:

1. **Σύμπλεγμα 1:** παθιασμένος, με αυτοπεποίθηση, θυελλώδης, ζωηρός, θορυβώδης
2. **Σύμπλεγμα 2:** διασκεδαστικός, χαρούμενος, γλυκός, αξιαγάπητος
3. **Σύμπλεγμα 3:** δριμύς, οδυνηρός, σκε-



Εικόνα 1: Κυκλική ταξινόμηση διάθεσης κατά Hevner

4. **Σύμπλεγμα 4:** χιουμοριστικός, ανόητος, ιδιότροπος, πνευματώδης, σαρκαστικός
5. **Σύμπλεγμα 5:** επιθετικός, φλογερός, έντονος, ενστικτώδης

Με σκοπό την ανάπτυξη ενός υπολογιστικού μοντέλου για την αυτόματη εξακρίβωση διάθεσης από σήμα μουσικής, υπάρχουν διάφορα ζητήματα που πρέπει να ληφθούν υπόψη:

1. Μια κοινή παρατήρηση στο πρόβλημα

της ανίχνευσης της φέρουσας διάθεσης ενός σήματος μουσικής, είναι το γεγονός ότι η αντίληψη της από τον ακροατή, είναι υποκειμενική και εξαρτάται από παράγοντες όπως, ο πολιτισμός, η εκπαίδευση του και η προσωπική του εμπειρία. Έτσι, για το ίδιο μουσικό κομμάτι, διαφορετικοί μουσικοί μπορούν να έχουν διαφορές στο τρόπο που θα το εκτελέσουν, ενώ διαφορετικά άτομα μπορεί να έχουν διαφορετικές αντιλήψεις για την διάθεση που το χαρακτηρίζει. Ως εκ τούτου, συχνά υπάρχει η άποψη ότι η διάθεση είναι μια πολύ υποκειμενική ποσότητα για να ανιχνευθεί. Ωστόσο, εκτεταμένη έρευνα στον χώρο αυτό έδειξε ότι συγκεκριμένα ακούσματα τα οποία φέρουν συγκεκριμένα σχέδια και δομές συνήθως έχουν εγγενή συναισθηματική έκφραση.

2. Υπάρχει συζήτηση, όπως είδαμε και σε προηγούμενη παράγραφο για το αν η συναισθηματική πληροφορία μπορεί να περιγραφεί με διακριτές κατηγορίες ή με συνεχείς ποσότητες. Δεν είναι επίσης ξεκάθαρο, ποια είναι τα βασικά συναισθήματα τα οποία μπορούν να εκφραστούν από την μουσική και να τα αντιληφθούν οι άνθρωποι. Οι ερευνητές χρησιμοποιούν επιθετικούς προσδιορισμούς για να περιγράψουν την διάθεση. Οι προσδιορισμοί αυτοί, χρησιμοποιούνται με μεγάλη ελευθερία με αποτέλεσμα η λίστα αυτών να είναι συνήθως τεράστια. Εκτεταμένη βασική έρευνα γίνεται σχετικά με την διάσταση (πλήθος) των βασικών συγκινησιακών καταστάσεων και αποτελεί τη βάση για την κατηγοριοποίηση της διάθεσης, που περιέχεται σε ένα μουσικό κομμάτι και δίνει βασικές ενδείξεις σχετικά με την μοντελοποίηση τους.
3. Το τρίτο ζήτημα, αφορά τα ακουστικά χαρακτηριστικά των μουσικών σημά-

των. Εκτός από κάποια χαρακτηριστικά, τα οποία μπορούν να εξαχθούν από την MIDI ή άλλων συμβολικών αναπαραστάσεων ενός μουσικού σήματος, υπάρχουν διάφορα ακουστικά χαρακτηριστικά τα οποία χρησιμοποιούνται για την κατηγοριοποίηση της πληροφορίας διάθεσης. Μολιταυτά, το μεγαλύτερο μέρος της μουσικής πληροφορίας υπάρχει σε μορφή κυματομορφής και δεν υπάρχει διαθέσιμο σύστημα αναπαράστασης το οποίο να την μετασχηματίσει σε συμβολική αναπαράσταση. Για αυτό το λόγο είναι απαραίτητη η αντιμετώπιση του προβλήματος κατηγοριοποίησης κάνοντας χρήση της ακουστικής αναπαράστασης του μουσικού σήματος. Μολιτονότι έχει γίνει εκτεταμένη έρευνα στην ανάπτυξη μουσικών και ακουστικών χαρακτηριστικών, παραμένει ακόμη ένα δύσκολο πρόβλημα η απευθείας αναπαράσταση του συναισθηματικού περιεχομένου μιας μουσικής ηχογράφησης. Κάποια από τα χαρακτηριστικά τα οποία έχουν χρησιμοποιηθεί για την μουσική ανάλυση σε επίπεδο κυματομορφής είναι οι:

- **mel-frequency cepstral παράμετροι** (MFCC),
- **η ενέργεια μικρού μήκους** (short-time energy, STE), και
- **ο ρυθμός μηδενισμού συνάρτησης** (Average Zero-Crossing Rate ή ZCR).

Κάποια χαρακτηριστικά τα οποία έχουν χρησιμοποιηθεί για την κατηγοριοποίηση συναισθημάτων σε σήματα ομιλίας, όπως η τονικότητα (pitch), δεν είναι εφικτή για την κατηγοριοποίηση διάθεσης στη μουσική. Κάποια όμως εξειδικευμένα ακουστικά χαρακτηριστικά του μουσικού σήματος μπορούν να χρησιμοποιηθούν για την κατηγοριοποίηση μουσικού σήματος σε είδος, όπως:

- **η χροιά** (timbre),

- ο ρυθμός και το μέτρο (rhythm and beat).

Αναγνώριση διάθεσης με βάση την εικόνα του προσώπου

Μια εικόνα για τον ανθρώπινο εγκέφαλο, αποτελείται από μια αποτύπωση αντικειμένων, ένα στιγμιότυπο. Όμως μια εικόνα για τον ηλεκτρονικό υπολογιστή είναι ένας (ή παραπάνω) πίνακας με δυαδικές τιμές, που αποτελούν τις φωτεινότητες και τα χρώματα που παράγουν την εικόνα αυτήν. Ένας Η/Υ δεν είναι σε θέση να αντιμετωπίσει μια εικόνα σημειολογικά, παρά μόνο σαν ένα σύνολο τιμών και δεδομένων. Όμως, πλέον στην επιστήμη της πληροφορικής έχουν αναπτυχθεί τρόποι και μεθοδολογίες, με τις οποίες ένας Η/Υ μπορεί να εκπαιδευτεί ώστε να αναγνωρίζει αντικείμενα μέσα σε μια εικόνα, και κατά επέκταση μέσα σε ένα βίντεο.

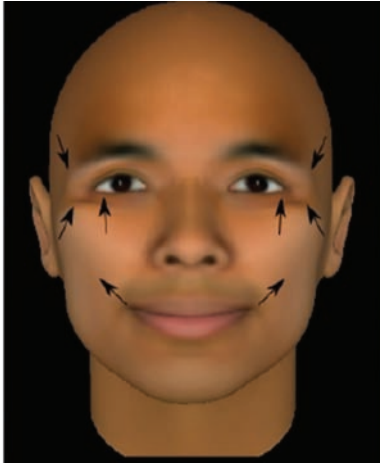
Η αποτύπωση της διάθεσης στην εικόνα και στο video, γίνεται βασικά με την κινησιολογία και την έκφραση του προσώπου. Η ταξινόμηση της κινησιολογίας είναι ένα αρκετά πιο δύσκολο πρόβλημα αυτόματης αναγνώρισης και προς το παρόν δεν θα το μελετήσουμε, αφού θα επικεντρωθούμε στην ταξινόμηση των εκφράσεων του προσώπου. Συνεπώς και σε αυτό το πλαίσιο, θα εξεταστούν τρόποι για την ανίχνευση ενός προσώπου μέσα σε μια εικόνα και στη συνέχεια τεχνικές για την αναγνώριση της διάθεσης όπως απεικονίζεται στο πρόσωπο.

Υπάρχουν πολλές τεχνικές για την εύρεση προσώπων μέσα στην εικόνα, και ποικίλουν σε πολυπλοκότητα αλλά και αποτελεσματικότητα. Ίσως ο πιο εύκολος τρόπος για να εντοπιστεί ένα πρόσωπο είναι να ελεγχθεί το παρασκήνιο της εικόνας, δηλαδή το φόντο να είναι απλό ώστε να είναι εύκολο με την κατάλληλη επεξεργασία να εντοπιστεί το πρόσωπο του ατόμου. Η εύρεση αυτή μπορεί να γίνει με βάση το χρώμα του ανθρώπινου δέρμα-

τος, το οποίο κυμαίνεται μέσα σε ένα γνωστό φάσμα χρωμάτων. Αυτό ενδέχεται να μην παρουσιάζει ικανοποιητικά αποτελέσματα για όλους τους ανθρώπους, καθώς μπορεί να υπάρξουν μεγάλες αποκλίσεις στο χρώμα του δέρματος. Μια άλλη μέθοδος που είναι περισσότερο αξιόπιστη είναι η ανίχνευση κίνησης μέσα σε ένα βίντεο. Αυτό πάλι προϋποθέτει τον έλεγχο του φόντου, καθώς θα πρέπει να είναι σταθερό και αμετάβλητο, όμως δεν εγγυάται και την ορθή ανίχνευση του ζητούμενου προσώπου, αφού κάθε στοιχείο το οποίο κινείται θα μπορεί να θεωρείται και το ζητούμενο στοιχείο. Μια άλλη αποτελεσματική τεχνική, βασίζεται στην κατάτμηση της εικόνας σε σχήματα και χρησιμοποιώντας τη γνώση του σχήματος του ανθρώπινου προσώπου και των χαρακτηριστικών του (μάτια, μύτη, κτλ), πετυχαίνεται ο τελικός εντοπισμός του προσώπου. Υπάρχουν πολλές ακόμα τεχνικές (όπως αλγόριθμοι εκπαίδευσης, νευρωνικά δίκτυα, κτλ) με τις οποίες μπορούμε να αναζητήσουμε πρόσωπα μέσα σε μια εικόνα, καθώς και συνδυασμοί μεθόδων για την βέλτιστη αποτελεσματικότητα τους.

Μετά τον εντοπισμό του προσώπου, το επόμενο βήμα είναι η αναγνώριση της έκφρασης. Η αξιολόγηση της έκφρασης, είναι πολλές φορές καθαρά υποκειμενική, παρόλα αυτά όμως υπάρχουν κάποιοι αντικειμενικοί κανόνες που ο άνθρωπος εγκέφαλος χρησιμοποιεί για να αναγνωρίσει τις βασικές ανθρώπινες εκφράσεις:

- **Χαρά:** Ένα συναίσθημα που σίγουρα το διακρίνουμε από το χαμόγελο. Στην εικόνα 2, παρατηρείται ότι στο πρόσωπο, όταν υπάρχει η έκφραση χαράς δεν είναι μόνο το στόμα αυτό που συμμετέχει, αλλά και τα μάτια. Οι άκρες του στόματος ορθώνονται προς τα πάνω, δίνοντας μας το χαμόγελο, ενώ τα ζυγωματικά συγκλίνουν προς τα μάτια, όπως και οι άκρες των φρυδιών.

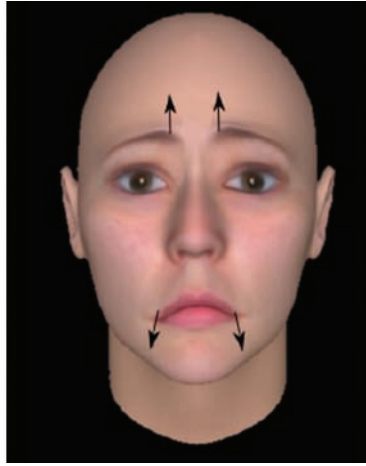


Εικόνα 2: Χαρακτηριστικά χαμόγελου

- **Θυμός:** Ανάλογα και την ένταση του θυμού γίνονται και πιο εμφανή τα σημάδια του στο πρόσωπο. Όπως διακρίνεται στην εικόνα 3, το κύριο γνώρισμα του συναισθήματος αυτού, είναι η σύγκλιση των φρυδιών προς την μύτη του προσώπου, και μεταξύ τους. Τα περισσότερα στοιχεία του προσώπου, κατά μέσω όρο είναι αμέτοχα, αφήνοντας σταθερό το στόμα σε φυσιολογική στάση.
- **Λύπη:** Η λύπη είναι ένα από τα συναισθήματα που είναι πολλές φορές δυσδιάκριτα. Το πρόσωπο του ανθρώπου

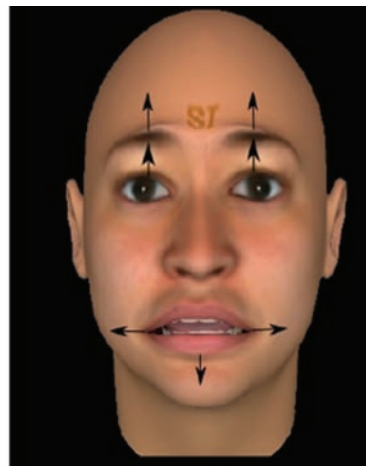


Εικόνα 3: Χαρακτηριστικά θυμού



Εικόνα 4: Χαρακτηριστικά λύπης

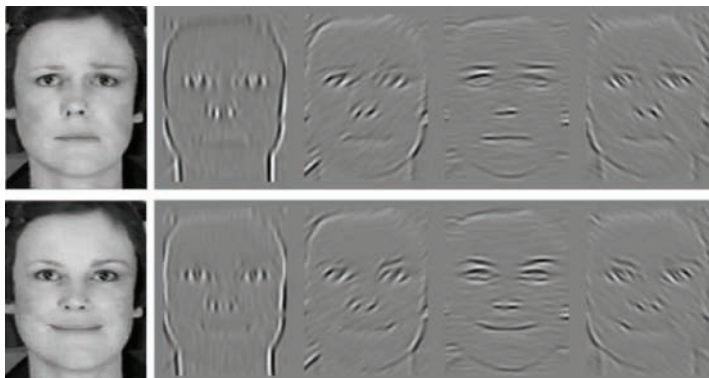
- μπορεί να μην εκφράζει πάντα με τον ίδιο τρόπο αυτό το συναίσθημα και πολλές φορές το μπερδεύουμε με την φυσική στάση του προσώπου. Όταν όμως είναι ξεκάθαρη η εκδήλωση του τότε, όπως παρατηρείτε και στην εικόνα 4, τα άκρα των χειλιών έχουν μια κλίση προς τα κάτω και τα φρύδια έχουν μια κλίση προς τα πάνω.
- **Φόβος:** Το συναίσθημα του το φόβου, είναι από τα συναισθήματα που συμμετέχουν περισσότερο τα στοιχεία του προσώπου. Υπάρχει συμμετοχή του στόματος, των ματιών και των φρυδιών του



Εικόνα 5: Χαρακτηριστικά φόβου

προσώπου. Τα φρύδια ανασπώνονται καθώς τα μάτια ανοίγουν περισσότερο από το συνηθισμένο. Το στόμα ανοίγει και τα χείλη τεντώνονται αριστερά και δεξιά του προσώπου.

Συνεπώς, το επόμενο βήμα είναι να απομονωθούν τα κύρια στοιχεία του προσώπου, δηλαδή το στόμα, τα μάτια και η μύτη. Μέσω αυτών των στοιχείων και αναλύοντας το σχήμα τους και τη θέση τους, θα μπορέσουμε να καταλάβουμε το συναίσθημα που εκφράζεται. Μια από τις μεθόδους που χρησιμοποιείται αρκετά για την εξαγωγή και παραμόρφωση χαρακτηριστικών του προσώπου, είναι ο μετασχη-



Εικόνα 6: Gabor μετασχηματισμοί

ματισμός Gabor, ο οποίος είναι ένας μετασχηματισμός Fourier με μια Gauss λειτουργία, προκειμένου να φιλτραριστούν οι τοπικές πληροφορίες της εικόνας. Η εικόνα 6, απεικονίζει μια εξαγωγή χαρακτηριστικών προσώπου με τη χρήση του Gabor. Στα αριστερά, παρατηρούμε το πώς είναι το πρόσωπο και στα δεξιά είναι τα αποτελέσματα Gabor, με διάφορα φίλτρα με διαφορετικές συχνότητες και κατευθύνσεις.

Ο Gabor μετασχηματισμός, τονίζει τις σημαντικότερες περιοχές του προσώπου, καθιστώντας πιο εύκολη της αποκοπή τους από το υπόλοιπο πρόσωπο. Έτσι, μέσα από μια διαδικασία φιλτραρίσματος της ει-

κόνας γίνεται εφικτή η ανεύρεση των στοιχείων του προσώπου, που είναι τα ζητούμενα. Πλέον, υπάρχουν απομονωμένα τα στοιχεία που χαρακτηρίζουν την έκφραση του προσώπου. Το θέμα είναι, πώς ο υπολογιστής θα μπορεί να καταλάβει την πληροφορία που υπάρχει, όπως ακριβώς την καταλαβαίνει και την ερμηνεύει και ο άνθρωπος. Η διαδικασία που ακολουθεί ο άνθρωπος είναι ότι αναλύει τη σκηνή, εντοπίζει το πρόσωπο και τα χαρακτηριστικά του, και επειδή ξέρει πως είναι το χαμόγελο, καταλαβαίνει ότι πρόκειται για ένα χαρούμενο πρόσωπο. Άρα, το μόνο που πρέπει να γίνει πλέον, είναι να εκπαιδευτεί και ο υπολογιστής για να γνωρίζει πως είναι το χαμόγελο και να ερμηνεύει αντιστοίχως τη σκηνή.

Σε αυτή τη περίπτωση, χρησιμοποιούνται συστήματα μάθησης μηχανής όπως τα νευρωνικά δίκτυα. Αρχικά, ορίζεται ένα αρχείο δεδομένων για εκπαίδευση του συστήματος, το οποίο θα περιέχει τα συναισθήματα που θέλουμε να αναγνωρίζει ο αλγόριθμος. Θα πρέπει να υπάρχει ένας ικανοποιητικός αριθμός εικόνων

που θα αντιπροσωπεύουν τα συναισθήματα. Έτσι, ο υπολογιστής κάνοντας την εκπαίδευση αναλύει διάφορες εικόνες και παίρνει χαρακτηριστικά για κάθε εικόνα, επικεντρώνοντας την προσοχή του στα βασικά στοιχεία της (στόμα, μάτια και μύτη). Αφού έχει τελειώσει την διαδικασία της εκπαίδευσής του, έχει μια γενική ιδέα του πώς μπορεί να είναι για παράδειγμα ένα χαμόγελο και συνδυάζει τα στοιχεία αυτά, για να συνδέσει τα συναισθήματα με κάποια γενικά γνωρίσματα. Όσο καλύτερη η βάση δεδομένων που του παρέχουμε για την εκπαίδευσή του, τόσο πιο αποτελεσματικός είναι στην αυτόματη διαδικασία εύρεσης συναισθημάτων προσώπου. Μια αρκετά αξιόπι-



Εικόνα 7: Διπλή ταξινόμηση διάθεσης

στη μεθοδολογία είναι η Linear Discriminant Analysis (LDA) η οποία χαρακτηρίζει και ξεχωρίζει δύο ή περισσότερες κλάσεις αντικειμένων. Σκοπός της μεθόδου, είναι η ανάπτυξη μιας σειράς διακριτικών συναρτήσεων, οι οποίες μεγιστοποιούν τη διακύμανση μεταξύ των κατηγοριών σε σχέση με την διακύμανση εντός των κατηγοριών. Μια καινούργια τεχνική που μπορεί χρησιμοποιηθεί να ταξινομήσει και με μεγαλύτερη ακρίβεια τα συναισθήματα του ανθρώπινου προσώπου είναι και η μέθοδος SIFT. Ο αλγόριθμος SIFT (Scale Invariant Feature Transform), χρησιμοποιείται για την ανίχνευση χαρακτηριστικών σημείων-κλειδιών (keypoints) σε εικόνες. Ο πρωταρχικός στόχος του SIFT, είναι η αναγνώριση προτύπων μορφών μέσα σε μια εικόνα. Ο SIFT, θεωρείται ιδιαίτερα εύρωστος και τα κλειδιά που αναγνωρίζει μπορούν να παραμένουν αναλλοίωτα σε μια σειρά από ασαφείς παραμορφώσεις, προσθήκες θορύβου και αλληλαγών φωτισμού, οδηγώντας σε περισσότερο αξιόλογα αποτελέσματα στην αναγνώριση εκφράσεων προσώπου.

Συνδυασμένη ανίχνευση διάθεσης

Με βάση τις παραπάνω αναλύσεις, είναι φανερό ότι ο συνδυασμός των αποτελεσμάτων μεθόδων ταξινόμησης της διάθεσης με βάση την μουσική επένδυση και με βάση την εικόνα/video θα μας προσφέρει μια πιο ασφαλή και πιο πετυχημένη εκτίμηση της διάθεσης (εικόνα 7). Το επόμενο βήμα, θα είναι η δημιουργία κοινών μεθόδων που θα χρησιμοποιούν ταυτόχρονα την οπτική και ακουστική πληροφορία για την εξαγωγή της ταξινόμησης της διάθεσης (εικόνα 8).



Συνδυασμένη ταξινόμηση διάθεσης

Εικόνα 8: Συνδυασμένη ταξινόμηση διάθεσης