



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
Επεξεργασία πρωτοομικών εικόνων

Ειρήνη Α. Κωστοπούλου

Επιβλέποντες: **Δημήτρης Μαρούλης**, Αναπληρωτής Καθηγητής
Ελένη Ζαχαρία, Διδάκτωρ

ΑΘΗΝΑ
ΔΕΚΕΜΒΡΙΟΣ 2009

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
Επεξεργασία πρωτομικών εικόνων

Ειρήνη Α. Κωστοπούλου

A.M.: ΠΙΒ006

ΕΠΙΒΛΕΠΟΝΤΕΣ: **Δημήτρης Μαρούλης**, Αναπληρωτής Καθηγητής
Ελένη Ζαχαρία, Διδάκτωρ

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: **Δημήτρης Μαρούλης**, Αναπληρωτής Καθηγητής ΕΚΠΑ
Θεοχάρης Θεοχάρης, Αναπληρωτής Καθηγητής ΕΚΠΑ
Μανώλης Σαγκριώτης, Αναπληρωτής Καθηγητής ΕΚΠΑ

ΔΕΚΕΜΒΡΙΟΣ 2009

ΠΕΡΙΛΗΨΗ

Η έρευνα στον τομέα της πρωτεομικής πραγματεύεται την συστηματική ανάλυση των πρωτεϊνών όπως αυτές εκφράζονται σε ένα δοσμένο κύτταρο, ιστό, ή βιολογικό σύστημα, σε μια δοσμένη χρονική στιγμή. Κατά την ανάλυση αυτή εντοπίζονται πιθανές αλλαγές στα επίπεδα των πρωτεϊνών - μεταξύ διαφορετικών βιολογικών καταστάσεων (φυσιολογικό, παθολογικό) - οι οποίες οφείλονται σε ποικίλες παθήσεις ή είναι αποτέλεσμα εξωγενών παραγόντων, όπως τοξικοί παράγοντες. Μια από τις βασικές απαιτήσεις στη μελέτη του πρωτεόματος είναι η ανάγκη εύρεσης αξιόπιστης τεχνικής διαχωρισμού η οποία είναι ικανή να διαχωρίσει πολύπλοκα πρωτεϊνικά μείγματα, ακόμα και χιλιάδων πρωτεϊνών σε ένα πείραμα. Η δισδιάστατη ηλεκτροφόρηση του πηκτώματος είναι μια ισχυρή, καθιερωμένη μέθοδος για τον διαχωρισμό χιλιάδων πρωτεϊνών μέσα σε ένα βιολογικό δείγμα. Με την δισδιάστατη ηλεκτροφόρηση του πηκτώματος οι πρωτεΐνες διαχωρίζονται σύμφωνα με το φορτίο τους και το μέγεθος τους σε ξεχωριστές κηλίδες σε ένα πήκτωμα. Η τοποθεσία όπου η πρωτεΐνη μεταναστεύει κατά την διάρκεια του διαχωρισμού είναι χαρακτηριστική για κάθε συγκεκριμένη πρωτεΐνη και το μέγεθος και η ένταση της κηλίδας εξαρτάται από την ποσότητα της πρωτεΐνης. Συνεπώς, μια πρωτεΐνη χαρακτηρίζεται από την θέση της πάνω στο πήκτωμα, και από γεωμετρικές παραμέτρους που περιγράφουν το σχήμα και τον όγκο της αντίστοιχης κηλίδας. Η πολυπλοκότητα των κηλίδων υπαγορεύει την χρήση δυνατών μεθόδων επεξεργασίας εικόνας ώστε να αναλυθούν τα πηκτώματα και να απομονωθούν οι πρωτεΐνες.

Στην παρούσα διπλωματική αναπτύχθηκε αλγόριθμος επεξεργασίας εικόνας για τον εντοπισμό των πρωτεϊνικών κηλίδων σε πηκτώματα δισδιάστατης ηλεκτροφόρησης. Η προτεινόμενη μέθοδος χρησιμοποιεί δισδιάστατα ιστογράμματα και μέθοδο κατάτμησης watershed για το διαχωρισμό των κηλίδων σε πρωτεΐνες. Η προσπάθεια υλοποίησης ενός τέτοιου αλγορίθμου κρίνεται απαραίτητη αφού τα διαθέσιμα λογισμικά επεξεργασίας εικόνων πρωτεομικής δεν παρέχουν ικανοποιητικά αποτελέσματα και συνήθως απαιτούν την παρέμβαση του χρήστη για την βελτιστοποίηση των αποτελεσμάτων. Η προτεινόμενη μέθοδος κατάτμησης κηλίδων εφαρμόστηκε σε συνθετικές και σε πραγματικές εικόνες και τα αποτελέσματα που προέκυψαν κρίνονται ικανοποιητικά.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Ανάλυση Εικόνας

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Βιοϊατρικά Δεδομένα, Εικόνα Πηκτώματος, Ιστόγραμμα, Κατωφλίωση, Κατάτμηση

ABSTRACT

Research in proteomics deals with the systematic analysis of proteins as they are expressed in a given cell, tissue, or biological system in a provisionally assigned time. In this analysis possible changes and alterations in the levels of proteins of different biological situations (normal, pathological) due to a variety of diseases or result of exogenous factors such as toxic agents are identified. One of the basic requirements in studying proteomics is the need to find a reliable separation technique that is able to separate complex protein mixtures, even thousands of proteins in one experiment. The two-dimensional gel electrophoresis is a strong, established method to separate thousands of proteins in a biological sample. The location where the protein migrates during the separation is typical for each particular protein and the size and intensity of stain depends on its amount. Therefore, a protein is characterized by its position on the gel, and by geometric parameters describing the shape and size of the spots. The complexity of the spots requires the use of specialized image processing methods in order to analyze gels containing isolated proteins.

The subject of this thesis was the development of an original algorithm for locating protein spots on two-dimensional gels electrophoresis images. The method is based on 2D-Histograms and watershed approach to segment the spots. The development of such algorithm is considered to be of critical importance since the available software packages provide unsatisfactory results and usually require user intervention in order to optimize their efficiency. The proposed segmentation method was evaluated using synthetic and real images and provided promising results.

THEMATIC AREA: Image Analysis

KEYWORDS: Biomedical Data, Gel image, Histogram, Thresholding, Segmentation